COMPOSITE INDEX DEVELOPMENT VIA DATA FUSION FOR SYSTEM MONITORING, DIAGNOSIS AND DECISION-MAKING

A Dissertation Thesis
Presented to
The Academic Faculty

by

Xinran Shi

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY in the
SCHOOL OF INDUSTRIAL AND SYSTEMS ENGINEERING

Georgia Institute of Technology
[May 2020]

COMPOSITE INDEX DEVELOPMENT VIA DATA FUSION FOR SYSTEM MONITORING, DIAGNOSIS AND DECISION-MAKING

Approved by:

Dr. Jianjun (Jan) Shi, Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Jionghua (Judy) Jin
Department of Industrial and Operations Engineering
The University of Michigan, Ann Arbor

Dr. Kamran Paynabar
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Ben Wang
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Chuck Zhang
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: [March 16, 2020]

To my beloved mother, sister and husband.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my advisor and mentor, Professor Jianjun Shi for his devoted supervision, continual encouragement, and insightful guidance. This thesis would not have been accomplished without his support. Working with and learning from him is such a memorable and invaluable experience for me. The joy and enthusiasm that he has for his research were contagious and motivational for me. His kindness, eagerness, and incredible patience have undoubtedly made a significant impact on the developments in my research and personal life.

My gratitude also goes to my thesis committee members, including Professor Jionghua (Judy) Jin, Professor Kamran Paynabar, Professor Ben Wang, and Professor Chuck Zhang. Their valuable suggestions and support not only help me make significant improvements to the dissertation but also have greatly aided me during my Ph.D. study.

I gratefully acknowledge the professors in academic society. In particular, Professor Jing Li for her kind support and assistance throughout my doctoral study. In addition, I would like to thank Dr. Jeffrey Hunt, Professor Zhiyong Liang, Dr. Andi Wang, Professor Xiaowei Yue, Mr. Zhen Zhong and all the colleagues in OG Technologies, Inc. for all the collaborations and opportunities they have provided me.

The members of the Dr. Shi group have contributed immensely to my personal and professional time at Georgia Institute of Technology. The group has been a source of friendships as well as good advice and collaboration. My Ph.D. study would have been incomplete without my colleagues and friends whom include but are not limited to: Mr.

Jialei Chen, Mr. Liexiao Ding, Ms. Yu Jin, Ms. Dan Li, Dr. Mohammed Nabhan, Ms. Tianxin Tang, Dr. Yuchen Wen, Ms. Chunxing Yin, and Dr. Yi Zhou. They have each provided me with both productive and constructive comments and accepted nothing less than excellence from me. In particular, I would like to give a special thanks to Mrs. Shi as for her tremendous encouragement and care.

I would wish to acknowledge the support from my family, with my mother especially, for supporting me throughout my life. For my sister, Ying, who is always optimistic and supportive for me. I am very fortunate to have their love and support. Lastly, I want to thank my husband, Hongyue Sun for his love, continual support and unflinching faith through this critical stage of my life. I shall be able to overcome any challenge in the rest of my life with his standing by me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Due to rapid advancements in sensing and computation technology, multiple types of sensors have been installed in advanced manufacturing systems. The sensor network automatically collects massive amounts of product and process information in real-time. Though each sensing data plays an important role in a given task, it is desirable to develop a composite index to provide an assessment of overall system performance based on multiple relevant sensing information.

This dissertation focuses on composite index development via data fusion for manufacturing system monitoring, diagnosis, and maintenance decision-making. In this study, we take efforts to develop methodologies that integrate advanced statistics and engineering domain knowledge to fuse complex sensing information and provide a reliable quantitative composite index. The composite index can reflect the system performance, and further be used in monitoring, diagnosis, and maintenance decision-making. The research results are illustrated with three manufacturing systems, including the carbon nanotube buckypaper (CNT buckypaper) fabrication process, hot-rolling process, and semiconductor manufacturing.

The thesis has five chapters. After the introduction, chapter 2 develops a composite index construction structure for in-situ quality monitoring and assessment of the CNT buckypaper manufacturing process. The composite index is developed by analyzing in-line Raman spectrum sensing data, which measures various quality characteristics of the CNT buckypaper. Yue *et al.* (2018) decompose Raman Spectra into fixed effects, normal effects, and defective effects. However, the extracted features cannot directly reflect the overall

quality characteristics of fabricated CNT buckypaper as they are groups of functional data. In this chapter, the proposed structure uses a weighted cross-correlation and maximum margin clustering to fuse the fixed effects into the inconsistency index. The weighted cross-correlation and variance analysis are used to fuse the normal effects into the uniformity index. Those individual quality features are then fused into a composite index to reflect the overall quality of carbon nanotube buckypaper. The proposed methods and effectiveness are demonstrated with real data collected from a fabrication lab.

In chapter 3, we proposed an automatic framework to identify process variables that are most influential to product quality via multiple permutation tests in the hot rolling process. We developed a MAnufacturing Data Engine (MADE) that achieves data fusion for initial data screening and analysis. In the MADE, a set of algorithms work in parallel to identify the process variables that are most relevant to the product quality variables. A case study is provided by using real production data to illustrate our proposed approach and to identify the most relevant process variables to the surface defects of rolling bars in a hot rolling process.

In chapter 4, we proposed a decision-making strategy for optimal proactive maintenance scheduling for cluster tools in a semiconductor manufacturing system. The proposed strategy is based on in-line sensing and system layout information. A sensor-driven AvaIlability-Degradation (AID) scheme is constructed by estimating the data-driven degradation trend of the chambers and integrating the degradation status to precisely estimate the expected availability of cluster tools.

Chapter 5 provides a summary of contributions and a discussion of future work.

CHAPTER 1. INTRODUCTION

1.1 Motivation

In a modern manufacturing system, numerous sensors are installed in each manufacturing device, equipment, or station for in-situ, real-time sensing of the process variables and product qualities. Though those sensing data provide rich information about the individual process variable status and the specific product quality issues, most of those sensing variables are monitored and analyzed individually for change detection and individual condition monitoring. It has been a challenging task on how to conduct data fusion to analyze those sensing data jointly, and to provide an effective performance measure for a given objective systematically.

The rapid development of sensor technology allows fast data generation and collection for industrial process monitoring, control, and decision-making support. As the size of the data increases exponentially, the amount of noisy data also grew significantly. For a given objective, the majority of value often comes from the minority of data. It is essential to squeeze the right information from massive data.

The modeling and analysis of those data often run into the following challenges: (1) The imbalance of data sets: massive normal operational data, but very limited abnormal operational data, which is especially true for a given type of conditions or failures; (2) lacking a unified model and strategy to make an informed decision which considering the complex data structure, data type, and acquisition rate.

Big data analytics for manufacturing system monitoring, diagnosis, and decision-making need a data-fusion driven composite index. The composite index helps system performance evaluation and improvement. A composite index is a particular designed math function that combines a group of variables or features to provide a useful statistical measure of overall system performance and to be used in effective decision-making for a given objective.

An excellent composite index should have the following characteristics: (1) be able to extract relevant information from massive data; (2) be effective in assisting decision-making for a given objective; (3) be scalable to high dimensional data; (4) be robust to noise and data uncertainties; (5) have clear interpretation capabilities with physical meaning; and (6) be easy to be calculated.

1.2 Research Objectives

In this thesis, we propose a systematic way to address composite index development problem, which includes the following studies:

(1) In a CNT buckypaper manufacturing process, Raman spectroscopy has been widely used to measure and evaluate fabricated CNT buckypaper materials. A Penalized Mixed-effects Decomposition (PMD) method has been developed for In-line Raman Spectroscopy data analysis that can decompose Raman Spectra into fixed effects, normal effects, and defective effects (Yue *et al.* 2018). This method can effectively extract features and the quality characteristics of continuous fabricated CNT buckypaper (e.g., consistency and uniformity properties), which provides the basis for the real-time system performance evaluation and quality assessment based on in-line Raman spectroscopy for CNT

buckypaper fabricating process. However, the outputs of this PMD data analysis lead to multiple dimensions about the CNT buckypaper performance, thus are not competent to determine the overall CNT buckypaper manufacturing system performance. A single meaningful composite quality index is desirable to provide an overall CNT buckypaper quality assessment. With this motivation, we develop indices to measure: (a) the between-sample differences (fixed effects) from the ideal sample through quantifying the weighted cross-correlation for profiles similarity and maximum margin clustering; and (b) the within-sample variation (normal effects) via the variance of weighted cross-correlation for profiles similarity. After quantifying these parameters, an overall composite index is proposed to represent the product quality of CNT buckypaper in a unified manner.

(2) In a hot-rolling process, multiple sensors are installed in multistage rolling stations to sensing different process variables. During the hot rolling processes, a set of process variables are collected from each station. These process variables are usually in the form of time-series data, representing the measurements taken from one end of the billet to the other end. Among those variables, the side and bottom temperature of a billet, the volume of cooling water shoot onto the billet (water flow), and the rolling speed of the billet are collected when a steel billet is passing through an intermediate stand. Process variables (e.g. speed) are also collected as billet enters and exits the non-twisting mill stand. In addition to those process sensing variables, an in-line optical sensing station, called HotEye® (Chang *et al.* 2009), is installed after the final rolling station. The HotEye® provides real-time, in-line measurements of surface quality of rolling bars and detects/classifies surface defects. Surface defects, such as seams and checkings, therefore, are measured and reported. Because a hot-rolling process has 50 to 80 rolling stations and

can be a miles-long production line, it is a challenging task to quickly identify the influential process variable once a surface defect is detected. In order to address this challenge, we propose an automatic data dependency identification approach based on multiple permutation tests. A composite index is proposed to reflect the influential rate of a process variable on the quality variable.

(3) The semiconductor industry is highly competitive in terms of quality, productivity, and cost. A cluster tool, commonly used in a wafer fabrication facility, is an automated robotic manufacturing system containing multiple computer-controlled process modules. A cluster tool processes multiple substrates via load lock chambers for loading substrates, a transfer chamber for transferring the wafers between chambers, and several process chambers for performing one or more processes. The cluster tool is often described as a small factory for its multiple functions, such as chemical vapor deposition, physical vapor deposition, and plasma etching. With the capital investments for cluster tools as high as billions of dollars, the equipment downtime would bring considerable loss of productivity and profit. As a result, maintenance is an important decision that is essential for improving equipment reliability and profit in a semiconductor fab. The proposed optimal proactive maintenance scheduling model considers the degradation of a chamber as well as its availability. Ultimately, we propose a proactive maintenance decision-making strategy based on in-line sensing information as well as the system layout information.

1.3 State-of-the-Art

Generally speaking, the composite index is a measure of changes in a representative group of individual data points (Babbie 2015). It is a compound measure that aggregates

multiple indicators. Many research areas have studied the concepts of the composite index. In health care, composite indicators are constructed for assessing the severity of disease (Best *et al.* 1976; Murray 1994; Jones *et al.* 2009) and evaluating the performance of the effectiveness of a physician (Liu *et al.* 2012). In economics, people have studied stock exchange composite index like a stock market index for decision-making by covering all common stocks listed on the stock exchange (De Bondt and Thaler 1995; Leigh *et al.* 2002; Lee *et al.* 2011).

An elaborate manufacturing system is equipped with lots of sensors to monitor various aspects of the system performance. Each sensor plays a vital role in showing certain aspects of the system performances. The data-rich environment has provided unprecedented opportunities for quality and productivity improvement (Ding *et al.* 2006). It is necessary to transfer the data-rich environment into an information-rich environment for inferencing the system behavior and optimal decision-making of a manufacturing system. The composite index is recognized as a useful tool for accomplishing the purpose of converting the data into information for the performance improvement of a manufacturing system (Liu *et al.* 2013). A composite index integrates all relevant sensing information into one decision-making index for a given objective. It is essential for quick and effective real-time decision-making. The Composite index is much easier to interpret than trying to find a common trend in multiple separate indicators (Nardo *et al.* 2005).

Depending on the level of implementation of a composite index, data fusion methods for constructing a composite index could be categorized into three classes (Volponi *et al.* 2004), which are sensor-level fusion, feature-level fusion, and decision-level fusion. Sensor-level fusion combines a set of correlated patterns of multiple sensors (Heger and

Pandit 2004; Liu *et al.* 2017) to monitor the performance of a system. The feature-level fusion combines extracted feature information of independent analysis methods (Goebel and Bonissone 2005), while the decision-level fusion combines the diagnostic actions, such as maintenance advisories, of a system to improve the overall system performance (Sun 2002; Fang *et al.* 2017; Yildirim *et al*. 2016a; Yildirim *et al.* 2016b).

1.4 Organization of the Thesis

The modern multistage manufacturing systems collect various types of data in each stage. How to infer from the collected data for a given objective becomes an essential problem. We need algorithms to link the multiple sensing information and construct one composite index. The manufacturing system will be improved via the composite index to make decisions for system monitoring, diagnosis, and maintenance decision-making. These are the common characteristics and common demands from all multistage manufacturing systems. As an example, in the CNT buckypaper fabrication process, there are multiple sensors and features deployed. Real-time quality assessment for Buckypaper becomes an essential problem. Along the hot rolling process, it also includes multiple stages and sensors. The efficient identification of the process variable, which influences the product quality, through data analysis becomes our objective. In the semiconductor manufacturing system, the cluster tool is an automated robotic manufacturing system that contains multiple computer-controlled process modules. Each module integrates with multiple sensors. Based on this rich data environment, making a saving-demonstratable proactive maintenance strategy becomes a big demand. This thesis studies in a systematic way of data fusion for composite index development.

Figure 1.1 outlines the structure of this thesis and also the relationship between the chapters.

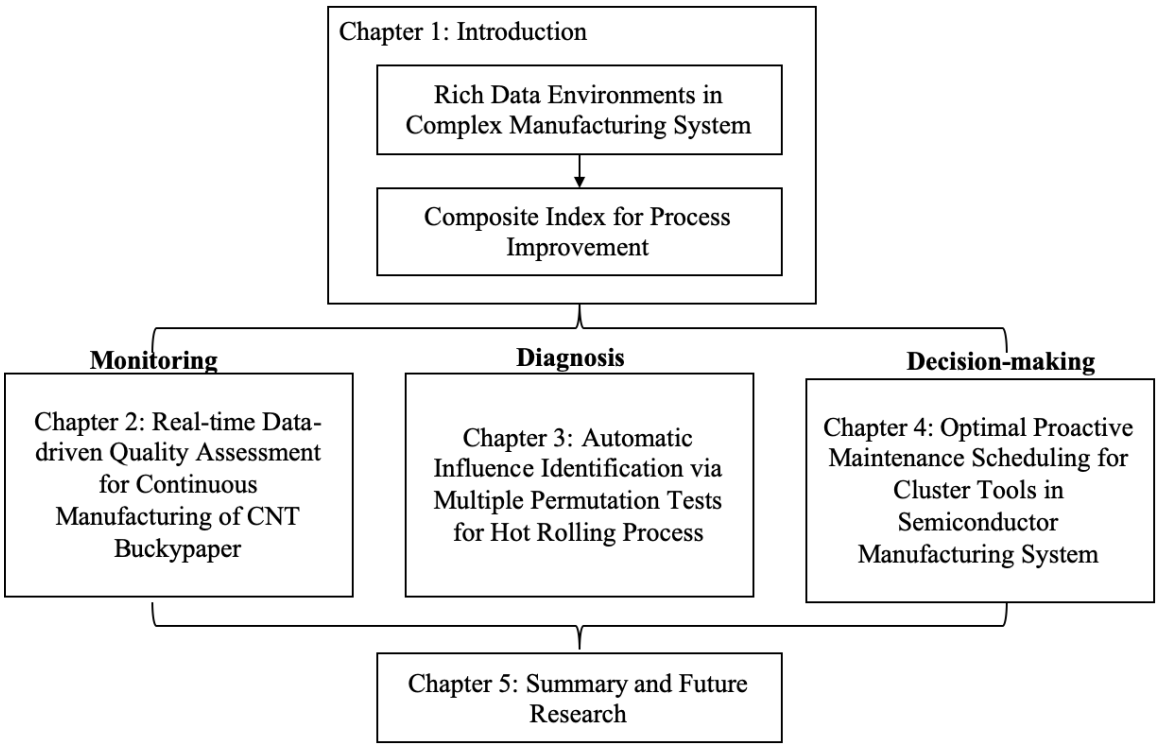


Figure 1.1 Outline of the Thesis

The thesis has five chapters. After Chapter 1 as an instruction, Chapter 2 develops a composite index for real-time data-driven quality assessment for the continuous manufacturing of carbon nanotube buckypaper. Carbon nanotube (CNT) thin sheet, or buckypaper, has shown great potential as a multifunctional platform material due to its desirable properties, including its lightweight nature, high mechanical properties, and good conductivity. A particular composite index is designed to reflect CNT buckypaper quality as a combination of consistency and uniformity properties. A case study indicates that our proposed quality assessment approach demarcates buckypaper quality in a unified manner.

In Chapter 3, we develop a framework to identify the influential process variable for a hot rolling process with multiple sensors installed in multistage rolling processes. The

proposed method does not require significant involvement of highly trained data scientists. A MAnufacturing Data Engine (MADE) is structured to achieve data fusion for initial data screening and analysis. It automatically identifies the influential process variables through multiple permutation tests. A set of algorithms work in parallel to identify which process variables influence the product quality variables. A case study is provided by using real production data to illustrate the proposed approach and to identify the most influential process variables on the product surface defects in a hot rolling process.

In Chapter 4, we propose an optimal proactive maintenance scheduling strategy for cluster tools in a semiconductor manufacturing system. Cluster tools are commonly used in semiconductor manufacturing and represent critical components of fab automation operations. In order to maximize productivity and throughput, a well-defined maintenance strategy is needed for semiconductor cluster tools. This chapter proposes an optimal proactive maintenance scheduling approach with a three-level hierarchy framework. The top-level model prepares a long-term planning horizon. The middle-level model estimates the reliability of cluster tools based on its degradation of the chambers and availability. The bottom- level is a Mixed-Integer Programming (MIP) model that considers the conditions of the cluster tools, and proactively provides an optimal maintenance schedule over the planning horizon. A case study is conducted via simulation to show that the proposed proactive maintenance strategy yields a better profit compared with current maintenance strategies.

In the end, Chapter 5 concludes the thesis and summarizes the original contributions. In addition, future research directions related to this thesis are discussed.

CHAPTER 2. REAL-TIME DATA-DRIVEN QUALITY ASSESSMENT FOR CONTINUOUS MANUFACTURING OF CARBON NANOTUBE BUCKYPAPER

2.1 Introduction

Buckypaper is a thin sheet made from an aggregate of carbon nanotubes, which could potentially provide high tensile strength, electrical and thermal conductivity, and optical properties (Krueger 2010). Researchers showed that carbon nanotube (CNT) buckypaper has excellent application potential as a superb multifunctional platform material with functionalities ranging from heavy-duty materials to electronic circuits protector to artificial muscles (Cha *et al.* 2013; Ajayan and Zhou 2001). However, the mass adoption and applications of CNT buckypaper have experienced significant bottlenecks because of the high cost in production and considerable uncertainty in quality. A systematic real-time quality assessment of high-performance buckypaper is urgently needed. The interests of the buckypaper characteristics include a specific type of multiwall carbon nanotube, geometric properties, width and diameter of the innermost wall, carbon unit cell ring size and connectivity, morphology, particle properties, and structural defects. Instead of studying these properties one-by-one, the macro perspective quality concerns of the CNT buckypaper include *consistency*, *uniformity*, and *defects*.

- *Consistency.* The degree of consistency indicates whether a gradual mean shift exits in the sequentially roll-to-roll fabrication process of CNT buckypaper.

- *Uniformity.* A sample is uniform if and only if the observations in the inspection area get similar features. The degree of uniformity reflects information such as the degree of alignment, the degree of functionalization, nanotube distribution, and dispersion.
- *Defects.* The within-sample defect information indicates whether there are defects in the CNT buckypaper. A specific band of Raman spectrum denotes corresponding defective information of the product.

Various measurement tools are applied for characterizing the properties of the CNT buckypaper, including scanning electron microscopy (SEM), transmission electron microscopy (TEM), fast Fourier transform (FFT) of high-resolution TEM (HRTEM), Raman spectroscopy, and Fourier transform infrared spectroscopy (FTIR) (Lehman *et al.* 2011). People usually use SEM for morphology and dimension measurements, and purity quantification (Trigueiro *et al.* 2007), and TEM and HRTEM for inner morphology measurements (including size, shape, purity, and disorder) (Li *et al.* 2008). FTIR spectroscopy can reflect the functionality of the product (Osswald *et al.* 2007). However, these techniques are not efficient nor applicable for real-time quality monitoring during the continuous nanomanufacturing process. Raman spectroscopy attracts broad interest for its potential on providing rich nanostructure information about the purity, defects, buckypaper functionality, and nanotube alignment. The offline characterization methods based on Raman spectroscopy have been widely used in batch-to-batch nanomanufacturing of CNT buckypaper (Gommans *et al.* 2000; Raravikar *et al.* 2002; Liu and Kumar 2003; Park *et al.* 2008). Although it has significant potential for quality monitoring (Févotte 2007; Alahbabi

et al. 2006; Abu-Absi *et al.* 2010), the real-time quality assessment based on inline Raman spectroscopy is not well studied yet.

As in-line Raman spectroscopy is a nondestructive testing and provides detailed nanostructure information within seconds, we use it to collect real-time data for quality assessment of the CNT buckypaper. Figure 2.1 shows the in-line Raman inspection for a 6-inch width roll-to-roll buckypaper production. For a sample zone in Figure 2.1 (a), one collects SEM pictures (Figure 2.1, b-d) from multiple sampling points for characterization, and inspects the corresponding Raman spectra (Figure 2.1, e-g) for real-time quality assessment. The Raman peak intensity ratio of D-band and G-band (I_D/I_G) determines the alignment degree of the samples (Févotte 2007; Cheng *et al.* 2009) and structural defects to graphitization or crystallinity ratio (Dresselhaus *et al.* 2010; Xu *et al.* 2008). However, the intensity ratio cannot tell the detailed information about the product.

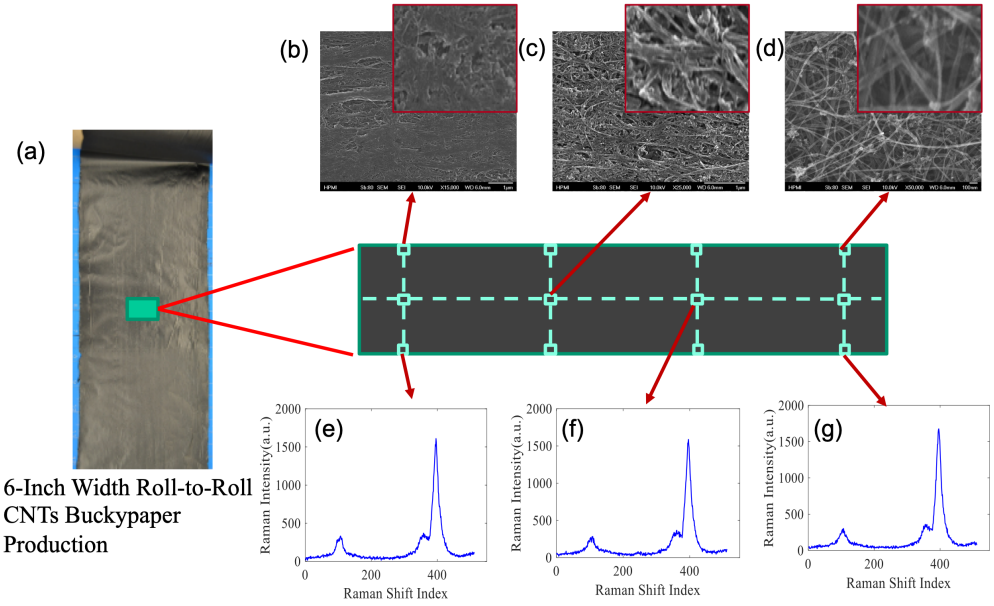


Figure 2.1 Inline Raman Spectroscopy Inspection for Buckypaper Production

From the SEM pictures (Figure 2.1, b-d), one could see that the fabrication of CNT buckypaper is not consistent as the degree of alignment become worse along the fabrication process, and the degree of uniformity within each sample is different. Hence, the intensity ratio cannot be used as process assessment and quality control guidance directly. Furthermore, the moving speed of the roll-to-roll buckypaper fabrication process is a massive challenge for high signal-to-noise ratio (SNR) signal acquisition of in-line Raman spectroscopy, which may result in significant uncertainties in the intensity ratio. Therefore, we need to develop a systematic quantification method for real-time quality assessment of high-performance CNT buckypaper.

Yue *et al.* (2017) used a generalized wavelet shrinkage method to increase the SNR of the in-line Raman spectra, which enables real-time quality control for CNT buckypaper manufacturing. Since the Raman spectra are collected from multiple channels and all the quality information, including consistency, uniformity, and defects, are mixed in the datasets. A data decomposition approach was developed, called wavelet-based penalized mixed-effects decomposition (PMD) (Yue *et al.* 2018), to obtain interpretable quality effects, i.e., (1) fixed effect that measures the fabrication consistency over time; (2) normal effects that reflect the uniformity of quality features within a sample; and (3) defective effects that indicate the existence and location of the defects in a sample area. A tensor mixed-effects model was also developed to separate fixed effects and random effects for high-dimensional arrays (Yue *et al.* 2020). Although the quality features decomposed from the PMD are interpretable and correspond to multiple quality characteristics, it cannot be used to conduct the real-time evaluation for the product quality of CNT buckypaper directly as the quality features decomposed from the PMD correspond to multiple high-dimensional

parameters. Jakubinek *et al.* (2019) proposed a quality assessment using X-ray photoelectron spectroscopy. Horne and Liang (2018) mentioned using in-line Raman spectroscopy for quality assessment. However, they cannot obtain a unified and quantitative index for quality assessment of CNT buckypaper. Moreover, the current practice of quality inspection, mainly based on operators' visual inspection, has three limitations: (1) subjective judgments by operators, (2) requirement of sophisticated training of operators, (3) slow reaction to the alert and lack of capability for real-time quality control. Therefore, a data-driven methodology is needed to perform a real-time quality assessment in a unified manner.

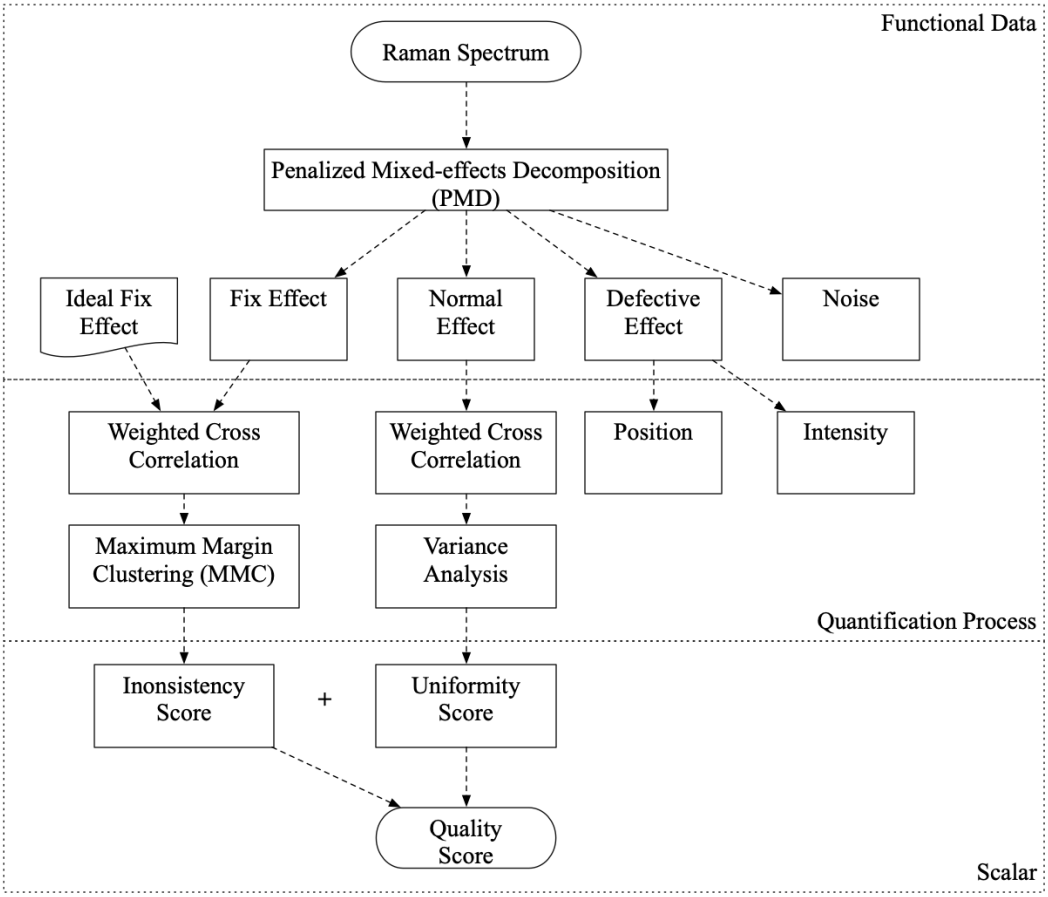


Figure 2.2 Flowchart for In-line Raman Spectroscopy Inspection of Buckypaper Production

The objective of this paper is to propose a standard real-time quality quantification methodology for the continuous manufacturing of CNT buckypaper, which directly reflects the quality information, such as impurity, alignment, functionalization, thickness, long-term consistency, and uniformity quickly and accurately. This real-time data-driven quality assessment methodology will provide quality control guidance to engineers for the CNT buckypaper industry.

As shown in Figure 2.2, after collecting in-line Raman spectra, we first apply the PMD algorithm (Yue *et al.* 2018) to extract fixed effects, normal effects, and defective effects from Raman spectra. Since one can easily interpret defective effects by the position and the intensity of the corresponding defects, we will focus on the study of the fixed effects and normal effects, which have no standard evaluation criterion. We use the proposed quality quantification approach to measure their between-sample differences (fixed effects) from the ideal sample and the within-sample variation (normal effects). After quantifying these parameters, an overall quality score is proposed to represent the product quality of CNT buckypaper in a unified manner. The obtained quality score has three key characteristics: (1) it should be distinguishable for CNT buckypaper samples with different quality levels, (2) it should be interpretable with corresponding physical features, (3) it should be easily obtainable from Raman spectroscopy inspection. The quality score provides guidance for operators to tackle process issues and improve the quality of CNT buckypaper. Based on the quality score, an operator could quickly determine if there is a process mean shift, local uniformity issue, or some defects. The stakeholders, such as customers, managers can realize intelligent decision making according to real-time data-driven quality assessment.

The remainder of this chapter is organized as follows. Section 2.2 illustrates a systematic CNT buckypaper quality assessment process. Section 2.3 presents a case study to demonstrate the implementation procedures. Finally, a summary is provided in Section 2.4.

2.2 Real-time Data-Driven Quality Assessment

2.2.1 Penalized Mixed-effects Decomposition (PMD) for In-line Raman Spectroscopy

The in-line Raman spectra are collected and classified into N groups according to a pre-designed maximin Latin Hypercube Design (Joseph and Hung 2008). A wavelet-based penalized mixed-effects decomposition (PMD) (Yue *et al.* 2018) is formulated for multichannel profile detection of in-line Raman spectroscopy as

$$\boldsymbol{y}_{ij} = \boldsymbol{\mu}_i + \boldsymbol{W}\boldsymbol{\theta}_{ij} + \boldsymbol{W}_{\boldsymbol{a}}\boldsymbol{\delta}_{ij} + \boldsymbol{e}_{ij} \tag{2.1}$$

where $\boldsymbol{y}_{ij}$ is a measurement profile with dimension n, corresponding to the j^{th} profile in the i^{th} sample; $\boldsymbol{\mu}_i$ denotes fixed effects in the i^{th} sample; $\boldsymbol{W}$ and $\boldsymbol{W}_{\boldsymbol{a}}$ are wavelet-based design matrices with dimension $n \times p$ and $n \times q$ for normal effects and defective effects respectively; $\boldsymbol{\theta}_{ij}$ and $\boldsymbol{\delta}_{ij}$ are coefficients vectors associated with normal effects and defective effects with regard to the j^{th} profile in the i^{th} sample. $\boldsymbol{e}_{ij}$ represents a signal-dependent noise vector for the j^{th} profile in the i^{th} sample. The three decomposed effects are represented as $\boldsymbol{\mu}_i, \boldsymbol{W}\boldsymbol{\theta}_{ij}$ and $\boldsymbol{W}_{\boldsymbol{a}}\boldsymbol{\delta}_{ij}$, which are (i) the fixed effects, $\boldsymbol{\mu}_i$, that reveals the fabrication consistency of the sample, i.e., long-term mean shift along the fabrication process, (ii) the normal effects, $\boldsymbol{W}\boldsymbol{\theta}_{ij}$, that quantify the uniformity of quality features in

the sample area, and (iii) the defective effects, $\boldsymbol{W_a \delta_{ij}}$, that shows the existence of specific sampling points with defective quality features.

However, the PMD has the following limitations when it is applied to real-time quality assessment for CNT buckypaper manufacturing process:

(1) It can decompose Raman Spectra into fixed effects, normal effects, and defective effects. However, a gap exists between the extracted features and the quality characteristics of continuous fabricated CNT buckypaper (e.g., consistency and uniformity properties).

(2) Although the PMD provides multiple dimensions to reflect CNT buckypaper performance, a single meaningful composite quality index is needed to reflect the overall CNT buckypaper quality.

To overcome those two limitations, we propose to quantify the fixed effects and the normal effects by using weighted cross-correlation to measure the similarity between profiles. The variance analysis is then applied to measure the deviation of the similarity of the normal effects.

2.2.2 Weighted Cross-correlation for Profiles Similarity Quantification

The fixed effects are driven from the mean vector of multiple profiles in each group of the sample that reflect the long-term mean shift of the fabricating process. This can be measured by the (dis)similarity between the sample's fixed effects and the ideal fixed effects. Similarly, the (dis)similarities among within-sample normal effects represent the degree of within-sample (dis)order.

There are multiple definitions of (dis)similarity measurement, whereas the best selection of a similarity measurement depends on specific domain knowledge. In the research area of the spectral library, an upcoming spectrum is searched among a known spectral library to find the optimal match. The similarity measure captures this match. The most common similarity measurements used in spectral library search are Euclidean, Mahalanobis, Pearson correlation coefficient absolute value, citiblock, cosine, and least square. The Euclidean similarity, citiblock, Mahalanobis, and correlation coefficient are classical point-to-point measurements, which are unable to deal with a minor shift and line broadening difference (Bodis 2007). The dissimilarity will increase significantly due to a small shift in peak positions if point-to-point measurements are used (de Gelder *et al.* 2001). However, the similarity measures which consider the neighborhoods can quickly capture the minor shifts. One popular similarity measurement which considers the neighborhoods is the weighted cross-correlation based generalized expression of similarity method (de Gelder *et al.* 2001). Inspired by de Gelder *et al.* (2001), we propose to apply weighted cross-correlation to quantify the long-term mean shift and the within-sample disorder that may happen in the CNT buckypaper fabrication process.

In a Euclidean vector space V, we define the profile as a continuous mapping $\boldsymbol{\beta_i}: [1, n] \rightarrow V$, where n is the dimension of the measurement profile $\boldsymbol{y}_{ij}$, $n \in \Re$ and $n \geq 1$. The ideal profile is $\boldsymbol{\beta}_0: [1, n] \rightarrow V$. Given two profiles $\boldsymbol{\beta}_0: [1, n] \rightarrow V$ and $\boldsymbol{\beta}_i: [1, n] \rightarrow V$, the cross-correlation function $c_{0i}(r)$ for pattern $\boldsymbol{\beta}_0(x)$ and $\boldsymbol{\beta}_i(x)$ is defined as:

$$c_{0i}(r) = \int \boldsymbol{\beta}_0(x)\boldsymbol{\beta}_i(x + r)dx \tag{2.2}$$

where r is the relative shift (lag) between those two functions, $\boldsymbol{\beta}_0(x)$ and $\boldsymbol{\beta}_i(x)$. The similarity between $\boldsymbol{\beta}_0(x)$ and $\boldsymbol{\beta}_i(x)$ is given by

$$\mathrm{S}_{0i} = \frac{\int (1 - \frac{|r|}{l}) \mathbb{I}_{\{|r|<l\}} \times c_{0i}(r) dr}{\sqrt{\int (1 - \frac{|r|}{l}) \mathbb{I}_{\{|r|<l\}} c_{00}(r) dr \int (1 - \frac{|r|}{l}) \mathbb{I}_{\{|r|<l\}} c_{ii}(r) dr}} \tag{2.3}$$

where l defines the width of the neighborhoods considered, $c_{00}(r)$ and $c_{ii}(r)$ are the auto-correlation functions that are defined in analogy to Equation (2.2), $\mathbb{I}_{\{|r|<l\}}$ is the indicator function that gets the value 1 for $|r| < l$, and the value 0 for $|r| \geq l$.

The dissimilarity between the pattern $\boldsymbol{\beta}_0(x)$ and $\boldsymbol{\beta}_i(x)$ is therefore given by:

$$D_i = (S_{00} + S_{ii} - 2S_{0i})/2 \tag{2.4}$$

where S_{00} and S_{ii} are the self-similarity of the pattern $\boldsymbol{\beta}_0(x)$ and $\boldsymbol{\beta}_i(x)$ respectively, and $S_{00} = S_{ii} = 1$.

To include the neighborhood into the calculation of the (dis)similarity, one should define the value of $\mathrm{r} \neq 0$. The dissimilarity criterion D_i will yield a value of 1 when the patterns of $\boldsymbol{\beta}_0(x)$ and $\boldsymbol{\beta}_i(x)$ are perfectly dissimilar, a value of 0 when patterns are identical, and a value between 0 and 1 for otherwise.

2.2.3 The Formulation for Inconsistency Index between Samples

The inconsistency of a process is the long-term mean shift that happened to the process. Since the fixed effect captures the long-term mean shift, the between samples' consistency assesses the dissimilarity changes of the fixed effects. We first adopt weighted

cross-correlation to measure the general dissimilarity. The two profiles in this case would be $\boldsymbol{\mu}_i(\mathrm{x})$ and $\boldsymbol{\mu}_0(\mathrm{x})$, where $i = 1, 2, \ldots N$ is the index of the samples, N is the total number of samples, 0 is the index of the ideal profile.

Besides, since the normalization in Equation (3.3) dilutes the mean shift of peak intensity, we introduce the maximum intensity difference ($\mathrm{d_i} = |\max(\boldsymbol{\mu}_0(\mathrm{x})) - \max(\boldsymbol{\mu}_i(\mathrm{x}))|$) to the consideration. For a sample i, let $\boldsymbol{z}_i^{\mathrm{T}} = (d_i, D_i) \in R^2$ be the row vectors of a collection of data points, arranged as the rows of the matrix $\boldsymbol{\mathcal{Z}} \in R^{N\times 2}$. Our main interest is to separate the data into consistent and non-consistent classes in a large margin classifier. Given data $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_N$, these data points would be assigned into two classes as $\eta_i \in \{-1, +1\}$, arranged as $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_N)^T$, when $\eta_i = -1$, the sample i is consistent with the others, when $\eta_i = 1$, the sample i is inconsistent with the others. In such a way, the separation between two classes is as wide as possible, which is known as unsupervised large margin method.

Unsupervised large margin methods, notably the maximum margin clustering (MMC) (Xu *et al.* 2005) is a popular clustering method that is motivated by the support vector machines (SVM). Without loss of generality, we assume the data set has been standardized as in the general procedure of MMC. Mathematically, the MMC approach aims at solving the following optimization problem (Xu *et al*. 2005):

$$\min_{\boldsymbol{\eta}} \min_{\boldsymbol{w},b} \left|\left|\boldsymbol{w}\right|\right|^2 + 2C\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{e} \tag{2.5}$$

$$\text{s. t.} \qquad \eta_i(\boldsymbol{w}^T\boldsymbol{z}_i + b) \geq 1 - \xi_i,\ \xi_i \geq 0 \tag{2.6}$$

$$\eta_i = \{\pm 1\}, -\ell \leq \boldsymbol{e}^T \boldsymbol{\eta} \leq \ell \tag{2.7}$$

where $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_N]^T$ is the vector of slack variables ($\xi_i, i = 1, \ldots, N$) for the errors, $C > 0$ is a regularization parameter and $\boldsymbol{e}$ is the vector of ones, and $\ell \geq 0$ is a constant controlling the class imbalance. This optimization problem can be solved by using the iterative approach (Zhang *et al.* 2009).

The distance from the data points to the optimal decision surface is used as the decision value in MMC. The confidence level of the probability for predicting true class increases when the decision value is large. The decision value, therefore, is an indicator of labeling consistent and inconsistent samples. To develop a single composite index for buckypaper consistency assessment, we adopt the decision value in MMC. The desired direction (e.g., small dissimilarity and maximum intensity difference) for buckypaper consistency assessment is already known. However, the signs of the decision value can be misleading to the engineers' intuition. To address this problem, we transfer the decision value by an arbitrary function, as described in Equation (2.6). The advantages of this transformation are: (1) when the shape parameter, the function is monotonic; (2) the interpretability of the index will be improved as the index would become non-negative in this study; (3) the sensitivity of the index will be improved as the Weibull cumulative distribution function will have a sharp increase at the boundary of shifting, and (4) the transform is invertible.

In our case, since $\boldsymbol{z}_i^{\mathrm{T}} \in R^2 \; and \; \boldsymbol{w^T} = (w_1, w_2) \quad \in R^2$, the decision surface is $\boldsymbol{w^T z}_i + b = 0$. The decision value is $\boldsymbol{\tau} = (\tau_1, \tau_2 \; \ldots, \tau_{\mathrm{N}})^{\mathrm{T}}$, where $\tau_i = \frac{|\boldsymbol{w^T z}_i + b|}{||\boldsymbol{w}||_2}$.

The inconsistency score is therefore defined as:

$$C_i = 1 - \mathrm{e}^{-\left(\frac{-\eta_i \tau_i - \min(-\boldsymbol{\eta} \circ \boldsymbol{\tau})}{\lambda}\right)^{\rho}} \tag{2.8}$$

where $\rho > 1$ is the shape parameter and $\lambda > 0$ is the scale parameter that needs further domain knowledge calibration according to the in-control data of a specific CNT buckypaper product. τ_i is the distance from a sample i to the decision surface, while η_i is the clusters that the sample belongs to (when $\eta_i = -1$, the sample i is consistent with the others, when $\eta_i = 1$, the sample i is inconsistent with the others). The elementwise matrix product of $\boldsymbol{\eta} \in R^N$ and $\boldsymbol{\tau} \in R^N$ is denoted by $\boldsymbol{\eta} \circ \boldsymbol{\tau}$, i.e., the Hadamard product.

The threshold of the inconsistency score is the value at $\tau_i = 0$, which means $\boldsymbol{z_i}$ is on the decision surface, and we further transform it to the inconsistency index space as:

$$\Delta = 1 - \mathrm{e}^{-\left(\frac{-\min(-\boldsymbol{\eta} \circ \boldsymbol{\tau})}{\lambda}\right)^{\rho}} \tag{2.9}$$

C_i reflects the changes in fabrication consistency due to the long-term process mean shift. The inconsistency index will be a value equal to 0 when the ideal fixed effects and the sample fixed effects are identical; otherwise, it would be a scaled number between 0 and 1. When $C_i \geq \Delta$, the sample i is inconsistent with other samples; otherwise, the sample is consistent with the others. The calculation steps are illustrated in *Algorithm 2.1*.

Algorithm 2.1 Inconsistency index procedure

Step1: Input the dissimilarity D_i as described in section 2.2.2. and the maximum intensity difference d_i.

Step 2: Initialize the labels $\boldsymbol{\eta}$ by simple clustering method.

Step 3: Fix $\boldsymbol{\eta}$ and train standard SVM model.

Step 4: Compute the $\boldsymbol{w}$ and b from the KKT conditions.

Step 5: Assign the labels as $\eta_i = sign(\boldsymbol{w}^T \boldsymbol{z}_i + b)$.

Step 6: Repeat *Step 3-5* until convergence.

Step 7: Return the labels $\boldsymbol{\eta}, \boldsymbol{w}$, and b.

Step 8: Compute the decision value $\tau_i = \frac{|\boldsymbol{w}^T \boldsymbol{z}_i + b|}{||\boldsymbol{w}||_2}$.

Step 9: Return the inconsistency score C_i from Equation (2.8).

2.2.4 The Formulation for Uniformity Index within Samples

The normal effects of sample i are denoted as $\boldsymbol{\theta}_{ij}(x)$, where $j = 1, \ldots, n$ is the index of normal effects within the sample. Two functions $\boldsymbol{\theta}_{ij}(x)$ and $\boldsymbol{\theta}_{ik}(x)$ are the normal effect functions in sample i, where x is the wavelength index. The similarity between $\boldsymbol{\theta}_{ij}(x)$ and $\boldsymbol{\theta}_{ik}(x)$ in sample i is defined as S_{jk}^i, while the cross-correlation, auto-correlation are defined as $c_{jk}^i(r)$, $c_{jj}^i(r)$ and $c_{kk}^i(r)$.

Within sample i, the similarity between observations j and k is known as S_{jk}^{i}. This similarity criterion will yield a value of 1 when the patterns of $\boldsymbol{\theta}_{ij}(x)$ and $\boldsymbol{\theta}_{ik}(x)$ are identical, and a value between 0 and 1 for other cases. However, this criterion cannot directly reflect the within-sample uniformity as there will have $n \times n$ similarity matrix for a sample.

The normal effects can provide us the information relevant to the degree of alignment, the degree of functionalization, nanotube distribution, and dispersion of the CNT buckypaper sample. The uniformity between the normal effects reveals relatively robust performance on alignment, functionalization, distribution, and dispersion. From the statistical perspective, the uniformity implies variability among mutual similarities of the observations within one sample. The uniformity within one sample i is then defined as:

$$U_i = \frac{\sum_{j=1}^{n} \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} \left(S_{jk}^{i} - \bar{S}_{j.}^{i} \right)}}{n} \tag{2.10}$$

where S_{jk}^{i} is the similarity between observations j and k in sample i, and $\bar{S}_{j.}^{i} = \sum_{k=1}^{n} S_{jk}^{i} / n$.

This index indicates the uniformity disorder of sample i due to within-sample random variations. The uniformity quantification criterion yields a scaled value from 0 to 1. A lower value of this index shows that the normal effects within the sample i tend to have better uniformity.

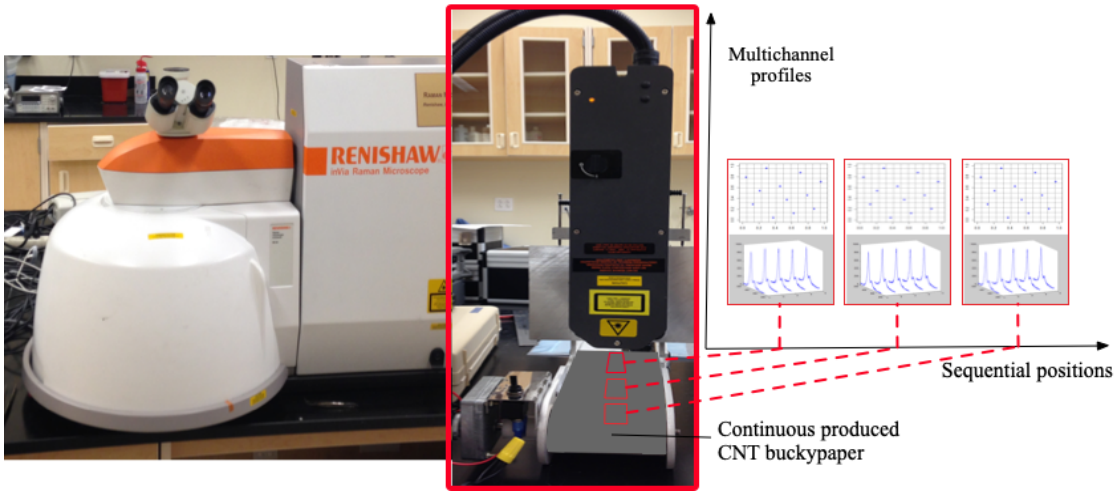
2.2.5 Overall Quality Quantification and Interpretation

To quickly check and rank the CNT buckypaper quality, a composed index is needed. Since the defective effects directly reflect the quality issue, we further use the consistency index (C_i), and the uniformity index (U_i) to compose the total quality index of the CNT buckypaper sample i:

$$\mathrm{Q}_i = \mathrm{W}_1 C_i + (1 - \mathrm{W}_1) U_i \tag{2.11}$$

where W_1 and $1 - W_1$ are the weights of the consistency and uniformity indices. These weights are chosen based on the significance level from engineering domain knowledge.

The total quality index yields 0 when the process is consistently, uniformly producing CNT buckypaper that are identical to the designed product. Otherwise, it would be a value between 0 and 1.



(a) *Renishaw™ Invia* Micro-Raman System

(b) Remote Optical Probe and Sample Roller

(c) In-line Measurement Data

Figure 2.3 Renishaw™ Invia Micro-Raman System with Custom-Designed Remote Optical Probe and Roller Sample Stage for In-Line Measurement.

2.3 Case Study

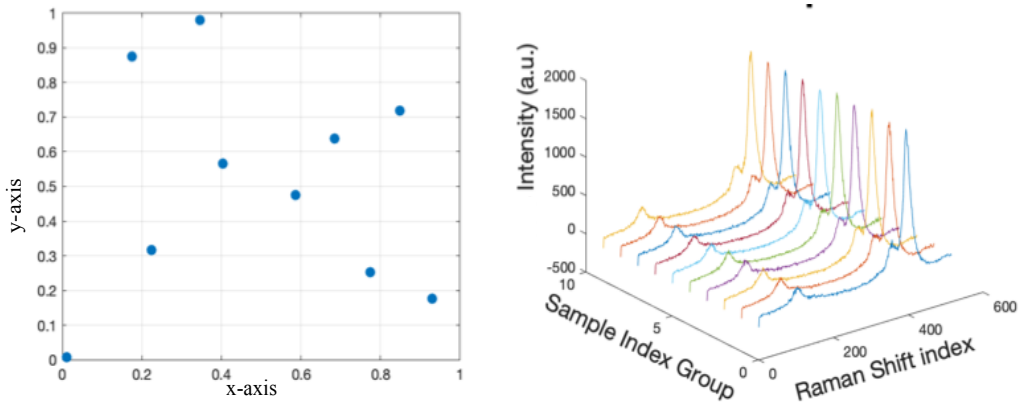
2.3.1 Experiment Preparation and Raman Spectra Interpretation

The fabrication detail of the CNT buckypaper with random alignment can be found in Park *et al.* (2008) and Gonnet *et al.* (2006). The CNT buckypapers' typical thickness in this experiment was 10 μm ~ 20 μm This thickness is measured by using Heidenhain-Metro incremental length gauge, and further confirmed by a SEM measurement Figure 2.3 shows the in-line Raman spectroscopy inspection system. Figure 2.3 (a) is the overlook of the Renishaw™ Invia Micro-Raman System. In the experimental set-up of this study, as shown in Figure 2.3 (b), we have a custom-designed remote optical probe and roller sample stage. For the remote probe, near-infrared (NIR) laser with a wavelength of 785 nm and a power of 150 mW were used to eliminate the effect of ambient lights. Low magnification lens was used to achieve a more considerable focus tolerance. Figure 2.3(c) shows that for each sample on the roller, the inspection system will measure the Raman spectra according to the per-determined design of experiment. The details about data collection will be discussed in the next subsection.

Raman spectra are used for detecting the quality information of the CNT buckypaper. It provides us with information relevant to potential quality issues such as impurity, the degree of chemical functionalization, and the alignment of CNTs in buckypaper. It is known that Radial Breathing Mode (< 300 cm^{-1}) is used to determine the diameter of single-walled carbon nanotube (SWCNT); D-band (1250 ~ 1400 cm^{-1}) is related to the disorder or molecular defects in the CNT structure, and D-band to G-band intensity ratio is a useful indicator of CNT quality of functionalization. In addition, polarized Raman provides angular dependence of the Raman intensity, and then the degree of CNT alignment can be estimated (Krueger 2010).

2.3.2 Data Collection

Figure 2.4 provides a further illustration of the design of the experiment and the collected Raman Spectra for the CNT buckypaper samples on the roller. Within each sample, we use a Maximum Latin Hypercube Design, as shown in Figure 2.4 (a), to pre-determine the positions of a certain number of Raman spectra that needs to be collected in a unit square. The design of the experiment of this kind has a good space-filling property and the first-dimension projection property. The Raman spectra, as shown in Figure 2.4 (b), located at the corresponding sample points, are collected for inspecting the quality of the CNT buckypaper. These Raman spectra are used to extract a quantitative index to represent the quality in the sample area.



(a) Maxmin Latin Hypercube Design in a Unit Square

(b) Collected Raman Spectra

Figure 2.4 The Design of Experiment and the Corresponding Raman Spectra

In our experiment, the Raman spectra are collected in the measurement zone with a rectangular shape with 120 × 120 micrometers. We collect ten samples, and within each of them, ten observation points are collected. These observations are tested sequentially and a Raman spectrum with 512 Raman shifts and intensities is collected for each measurement point. All the Raman spectra are collected based on a piece of single-wall CNT buckypaper. In the Renishaw Invia Micro-Raman System, Raman Microscopy with

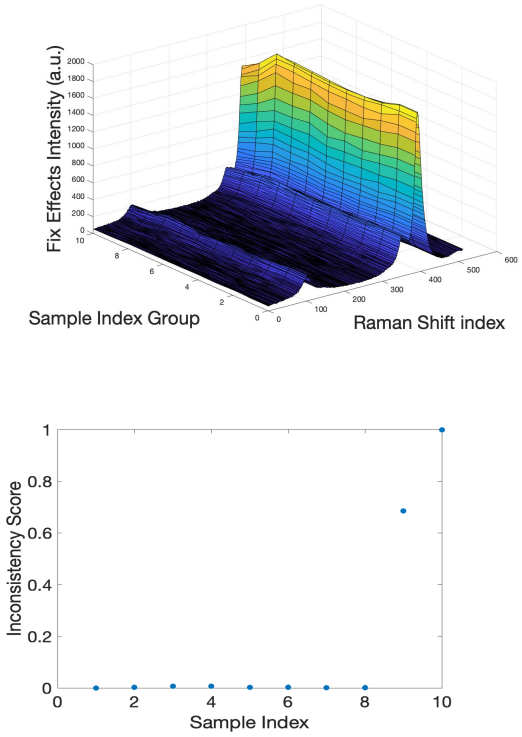
785 nm laser source and 0.5-second exposure time for each measurement point are conducted.

If each representative sample can be regarded as a sensor channel to collect Raman spectra, the process modeling and detection for the CNT buckypaper fabrication process can be formulated as a multichannel profile modeling problem along the sequential position of the CNT buckypaper. Thus, we use the PMD to process the collected Raman spectra.

2.3.3 Results and Discussions

After processing by the PMD, the Raman spectra are decomposed into fixed effects, normal effects, defective effects, and signal dependence noise. Figure 2.5 compares the fixed effect with the corresponding inconsistency index. Figure 2.5 (a) shows the comparison between ideal Raman spectra's fixed effect and the samples' fixed effects. The 0^{th} sample is the ideal fixed effect, while the 1^{st} to 10^{th} samples are the fixed effects decomposed from the real data, as shown in Figure 2.5 (b). The shape parameter and the scale parameter are selected to be 5 and 2, respectively, in Equation (2.6). Via empirical study based on in-control testing datasets, we find the inconsistency function has a sharp slope when the shape parameter is equal to 5. When the scale parameter is equal to 2, the minimum value and the maximum value of the slope are at the margin which separates the samples from the consistent and inconsistent groups.

Table 2.1 compares the long-term mean shift detection of the dissimilarity, maximum intensity difference, and the proposed inconsistency. By using dissimilarity alone, Exponentially Weighted Moving Average (EWMA) detects the 2nd sample beyond the limits. The Cumulative Sum (CUSUM) control chart for the maximum intensity difference detects the 10th sample beyond the limit. The proposed method detects that the 9th and 10th samples have mean shifts, which is identical to the underlying true observation shown in Figure 2.5 (a). Overall, the samples have a long-term mean shift comparing with the ideal fixed effects, especially from the 9th and 10th samples.

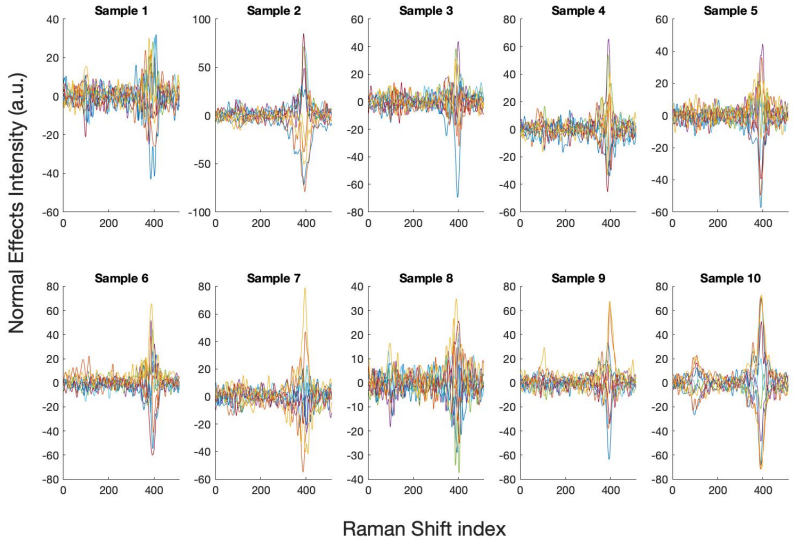


(b) Relative Inconsistency Score

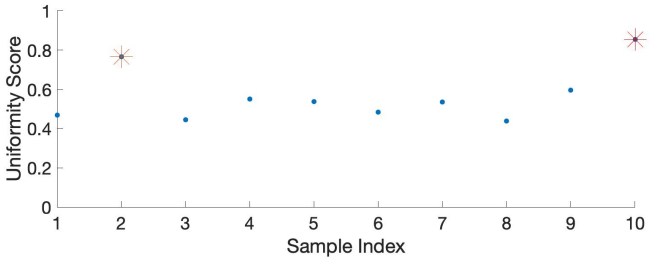
Figure 2.5 Comparison between Fixed Effects and the Corresponding Inconsistency Score

The inconsistency index is implemented to capture the long-term mean shift of the real data sample, as shown in Figure 2.5 (b). The consistency rank of the samples from the

good to the bad is the order of the 1st, 8th, 7th, 6th, 2nd, 5th, 4th, 3rd, 9th, and 10th samples. The Raman intensity after the 9th sample tends to be larger than the previous samples and the ideal sample. This change may result from the measurement equipment as the focus depth changes due to the sample local deformation. Another possible reason could be process changes. Some unknown process changes shift the process mean accidentally so that the 9th sample starts to be inconsistent with the other samples.



(a) Normal Effects of the Real Data



(b) Corresponding Uniformity Score

Figure 2.6 Comparison between Normal Effects and the Corresponding Uniformity Score

Table 2.1 Long-Term Mean Shift Detection Comparison Among Dissimilarity, Maximum Intensity Difference and Inconsistency (Bold and Underlined Number are Beyond Limits)

	Dissimilarity						***Maximum Intensity Difference***						***Inconsistency****	
	CUSUM ($h = 5$)			EWMA ($\lambda = 0.2$)			CUSUM ($h = 5$)			EWMA ($\lambda = 0.2$)			Proposed Method	
Sample Index	D	C_i^+	C_i^-	z_i	LCL	UCL	d	C_i^+	C_i^-	z_i	LCL	UCL	C_i	UCL
1	0	0	-1.85671	0.000104	9.69E-05	0.000163	0	0	-1.8325	89.0412	82.67089	139.9321	0	0.2899
2	0.00002	0	-3.35086	**8.72E-05**	8.76E-05	0.000172	65.55	0	-2.2913	84.34296	74.63642	147.9666	0.00197	0.2899
3	0.00008	0	-3.75729	8.58E-05	8.26E-05	0.000177	89.96	0	-2.23854	85.46637	70.3127	152.2903	0.00691	0.2899
4	0.00016	0.043857	-2.71343	0.000101	7.97E-05	0.00018	98.285	0	-2.01132	88.03009	67.77031	154.8327	0.00691	0.2899
5	0.00016	0.087714	-1.66957	0.000112	7.79E-05	0.000182	84.605	0	-2.07079	87.34508	66.21836	156.3846	0.00277	0.2899
6	0.00008	0	-2.076	0.000106	7.68E-05	0.000183	69.675	0	-2.44314	83.81106	65.25256	157.3504	0.00166	0.2899
7	0.00022	1.131571	0	0.000129	7.61E-05	0.000184	54.745	0	-3.12837	77.99785	64.64493	157.9581	0.00006	0.2899
8	0.0002	1.900571	0	0.000143	7.56E-05	0.000184	40.205	0	-4.11831	70.43928	64.26017	158.3428	0.00001	0.2899
9	0.00014	1.581857	0	0.000142	7.53E-05	0.000185	241.72	2.233127	-0.88518	104.6954	64.01556	158.5874	**0.68565**	0.2899
10	0.00024	3.076	0	0.000162	7.52E-05	0.000185	368.27	**7.11831**	0	157.4103	63.85968	158.7433	**0.99989**	0.2899

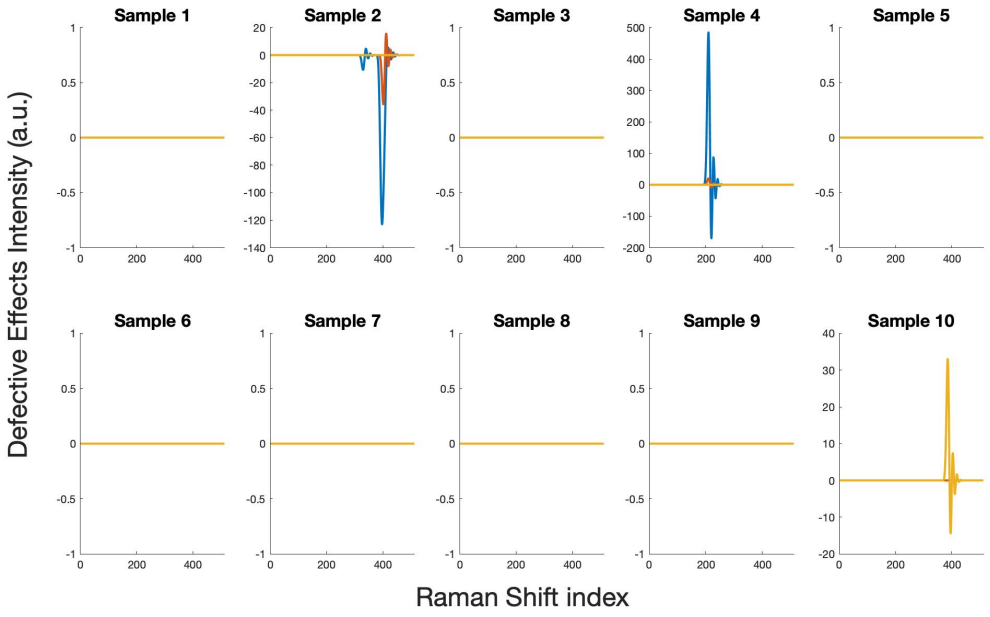


Figure 2.7 The Defective Effects of the Real Data

The normal effects of the real data reflect the degree of alignment, the degree of functionalization, nanotube distribution, and dispersion of the sample at the profile level. As shown in Figure 2.6 (a), it is hard to observe the entropy of the normal effects within the sample. Therefore, the uniformity index is calculated, and the corresponding uniformity index for the samples is shown in Figure 2.6 (b). The width of the neighborhood considered l is 2. The rank of the samples' uniformity from the good to bad is the 8^{th}, 3^{rd}, 1^{st}, 6^{th}, 7^{th}, 5^{th}, 4^{th}, 9^{th}, 2^{nd}, 10^{th} sample. The uniformity of the 2^{nd} and 10^{th} samples reflect the product quality issues, such as the degree of functionalization, the degree of alignment changes, nanotube distribution, and dispersion.

Figure 2.7 shows the decomposed defective effects of the read data. After the PMD extracts defective effects from the Raman spectra, one can observe that the observations #1 and #2 in the 2^{nd} sample, observations #8 and #9 in the 4^{th} sample, and observations #9 in the 10^{th} sample are defective. Based on the Raman shift, one can find that the defective effects of the 2nd and 10th samples occur at the G-band. The defects are due to the impurity of raw material, different degrees of functionalization, different alignments of carbon nanotubes, or bad nanotube dispersion. These quality issues can also be reflected in the uniformity quantification. The defective effect of the 4th sample is located between the D-band and the G-band, and it might result from some measurement errors, such as external light disturbance.

The defective effects of the 2nd and 10th samples occur in the G-band. The defects are due to the impurity of raw material, different degrees of functionalization, different alignments of carbon nanotubes, or bad nanotube dispersion. These quality issues can also be reflected in the uniformity quantification. The defective effect of the 4th sample is located between the D-band and the G-band, and it might result from some measurement errors, such as external light disturbance. One can keep the 4th sample in mind, and we will discuss this measurement error later. Figure 2.8 shows the overall quality performance of those ten samples ($w_1 = 0.3$). The overall quality rank of the samples from the good to bad is the 8th, 3rd, 1st, 6th, 7th, 5th, 4th, 2nd, 9th, 10th sample.

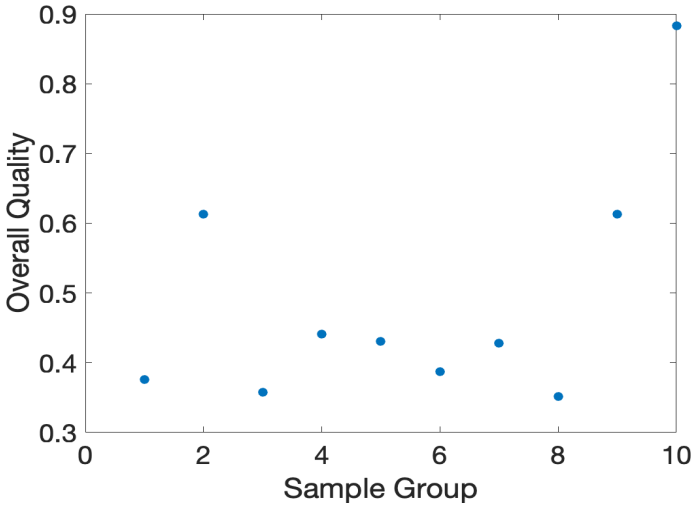


Figure 2.8 The Overall Quality of the Real Data

Although the overall quality index does not consider the defective effect, one can find that the samples with defects will be identified as not qualified samples by the proposed index. From the overall quality values, we could see that the 4th sample has a relatively smaller value than other defective samples. The quality of the 4th sample is comparable with samples without defective effects. In (Yue *et al.* 2018), the authors suspect the observed defective phenomenon on the 4th sample results from some

measurement errors, such as external light. Our overall quality index helps to verify that measurement errors cause the defective effect of the 4th sample.

Table 2.2 The Overall Quality of Three Buckypaper Materials

Material \ Sample ID	1	2	3	4	5
Raw SWCNT	0.40781	**0.60958**	0.40624	0.38501	**0.59167**
Acid SWCNT	0.3912	**0.63058**	0.42267	0.48447	**0.59869**
Functionalized SWCNT	0.40339	**0.57911**	0.40266	**0.57829**	0.43248

We further apply the proposed quality assessment method to three buckypaper materials. They are raw SWCNT, SWCNT after acid, and SWCNT after functionalization. Table 2.2 summarizes the overall quality score for all the samples of the materials. The threshold under the case study setting is 0.5. Therefore, the samples 2 and 5 of raw SWCNT and acid SWCNT have relatively low quality, while the samples 2 and 4 of functionalized SWCNT have relatively poor quality as their overall quality scores are beyond 0.5. This result is matching up with the conclusion by the experienced engineer.

2.4 Conclusion

In the continuous CNT buckypaper manufacturing process, a complicated profile data, called in-line Raman spectroscopy, is used to collect Raman spectra for the CNT buckypaper quality monitoring. The Raman spectra fuse affluent information that includes quality consistency, local uniformity, and within-sample defects. The PMD method enables us to extract fixed effects, normal effects, and defective effects from the Raman spectra. Although multiple quality features are decomposed from the Raman spectra, these features cannot be used to evaluate the real-time fabrication quality of CNT buckypaper

directly. Current practice relies on heuristic methods based on these quality features, which has specific limitations, including (i) subjective judgments by operators, (ii) requirement of sophisticated training of operators, and (iii) slow reaction to the process changes. It is crucial to develop novel quality assessment indices for the system to evaluate the product quality in a real-time manner automatically.

The main contribution of this chapter is to propose a new real-time quality assessment index to access the quality characteristics of samples based on in-line Raman spectra in a continuous CNT buckypaper manufacturing process. The proposed quality assessment indices quantify the CNT buckypaper quality from fixed effects and normal effects. The inconsistency index that derived from fixed effects reveals the long-term mean shift of the process, while the uniformity index that originates in normal effects reflects the within-sample uniformity. The overall quality index considers both uniformity and consistency to evaluate the quality of the CNT buckypaper. All these three indices yield from zero to one to show the corresponding quality characteristics from good condition to bad condition. In the case study, the proposed assessment approach is applied to distinguish the quality performance of the different CNT buckypaper samples. The proposed indices successfully identify the long-term mean shift that occurs in the process as well as the samples with the substantial within-sample disorder. Also, our proposed approach can provide quantitative quality indices for single-walled CNT buckypaper after acid processing or functionalization. The quality assessment results are consistent with evaluations from the experienced engineers. The interpretation based on the proposed quantification indices is clear to the corresponding physical features. It is not only obtainable from Raman spectroscopy inspection but also controllable for operators to

tackle process issues. By applying the proposed assessment method, the quality of CNT buckypaper could be quantified from the consistency, uniformity, and defective perspectives.

In future work, the correlation between the proposed indices and the control variables of the CNT buckypaper fabricating process will be studied. We will explore systematic in-line quality assessment and process improvement methodologies.

CHAPTER 3. AUTOMATIC DATA INFLUENCE IDENTIFICATION VIA MULTIPLE PERMUTATION TESTS IN HOT ROLLING PROCESS

3.1 Introduction

Enabled by advanced sensing technology, a hot rolling process has been implemented with numerous sensors to collect product quality and process variables in real-time from multiple rolling stations. Examples of those sensing data include rolling bar temperature, rolling speed, and the flow rate of cooling water in each rolling station. These data provide unprecedented opportunities to understand the manufacturing process variables that influence the surface defects on the products like checking and seams, which are critical quality concerns for rolling bars as they may lead to catastrophic failures once the product is in use.

However, the task of the data analytics for identifying the process variables that are related to or dependent on the surface defects is typically time-consuming and requires the involvement of well-trained data scientists. This requirement significantly constraints and limits the practical usage of those real-time sensing data for quality improvement. Therefore, it is desirable to develop a data analytics tool for the hot rolling process, which automatically reads the process and quality data and finds the dependence between product quality defects and specific process variables with minimum data scientist involvement. In other words, a useful and automatic data analytics tool for identifying potential causes of process conditions on quality issues without a data scientist is an urgent need.

Multi-sensor data fusion systems have been widely used to extract useful information from manufacturing process data. For example, data fusion based on the state-space model has been proposed and has been extensively studied (Shi and Zhou 2009). However, with the development of advanced sensing technology, the data generated from each stage of a manufacturing process are of the growing volume and getting ever more complex. Consequently, understanding the relationship between the product quality variables and process variables becomes a challenging task. Due to such complexity, understanding the manufacturing system and developing appropriate analytical tools to identify the process data that affects the quality variable require tremendous effort. The challenges of data analysis for hot rolling processes are illustrated in detail from the previous research (Jin *et al.* 2008), which can be summarized as three aspects:

(1) The data streams generated from each rolling station are multivariate and non-stationary. Because the data from all steps represent different physical measurements, these functional measurements take distinct shapes. For example, some speed measurements are staircase functions, some temperature measurements are continuous waveform functions, and some water flow rate measurements are of arbitrary shape. For this reason, distinct analysis methods should be applied to the data obtained from each individual stage.

(2) The product quality variables, as well as process variables from a rolling process, may fall into multiple clusters given different rolling conditions like production cycle, maintenance schedule, rolling material, and product specifications. The signals of products within the different groups are significantly different from the other, while the signals within each group are subject to less variation. An automatic grouping, or

clustering, is needed as part of the pre-processing of the data before further data analysis.

(3) The dependent relationship between the process data and the product quality variables takes complex forms. As introduced in (Jin *et al.* 2008), it is observed that the defects in the hot rolling process have a dependent relationship to a small number of features extracted from the process. Besides, multicollinearity is common in the raw dataset. The typical dependent process variable interpreted from the statistical significance of coefficients is not reliable, and sometimes misleading. This is because other independent process variables influence the marginal contribution of an independent process variable. Therefore, it is critical to understand whether a process feature generated from a process variable affects the appearance of defects on the product surface.

In the literature, several data analytics methods have been proposed on surface defects detection (Pan *et al.* 2009; Li *et al.* 2007). However, an automatic, systematic tool, which identifies the process variables and associated manufacturing steps that influence the quality variables on products through analyzing the data collected from the rolling processes, is demanded.

To address those challenges, we present a generic data analytic approach called ManufAcaturing Data Engine (MADE). MADE takes in the product quality data and the process variable data with various patterns, and automatically identifies and prioritizes the dependency between the process variables and the product defects. It provides meaningful

information on identifying the process variables influencing the quality variable for quality improvements.

The MADE framework consists of six steps, including data preprocessing, feature extraction, clustering and outlier detection, multiple permutation tests, decision-level fusion, and influential variable identification. This framework effectively learns from the data collected from the hot rolling process and quickly identifies process variables that influence the product defects.

The contributions of this chapter are summarized as follows. First, an integrated framework is proposed to find the connection between the process variable and the quality variable on the products in hot rolling processes. Second, we propose a new method that identifies the dependency between the features obtained from each process variable and the product defect using multiple types of statistical models via multiple permutation tests and the aggregated dependency indicator. With this method, we identify the process variables that influence the quality variables of the products by integrating several distinct statistical methods available in the literature.

The remainder of the chapter is organized as follows. The related background of the hot rolling process and its data acquisition are introduced in Section 3.2, and the existing approaches for holistic manufacturing data analysis are reviewed as well. Then, Section 3.3 describes the overall analytical procedure of the MADE. Section 3.4 introduces our proposed methods for identifying the process features that influence product quality. Section 3.5 presents a simulation study to evaluate the performance of the proposed method, and a case study by applying the proposed analysis procedure to a real production

data sets obtained from a rolling process. Finally, discussions and conclusions were given in Section 3.6.

3.2 Background of the Hot Rolling Process

In this section, we first briefly introduce the hot rolling process, and then discuss the challenges of using conventional approaches in manufacturing data analytics.

3.2.1 Background of the Hot Rolling Process

Hot rolling as the key steel-making process converts cast or semi-finished steel into finished products. A typical hot rolling process usually includes a melting division and a hot rolling division, as shown in Figure 3.1. The melting division is a continuous casting process. It melts scrapped metals and solidifies the molten steel into semi-finished steel billet. As for the hot rolling division, it squeezes the steel billet with a sequence of multiple stands that each consists of several rolls. There two types of stands: intermediate stand and

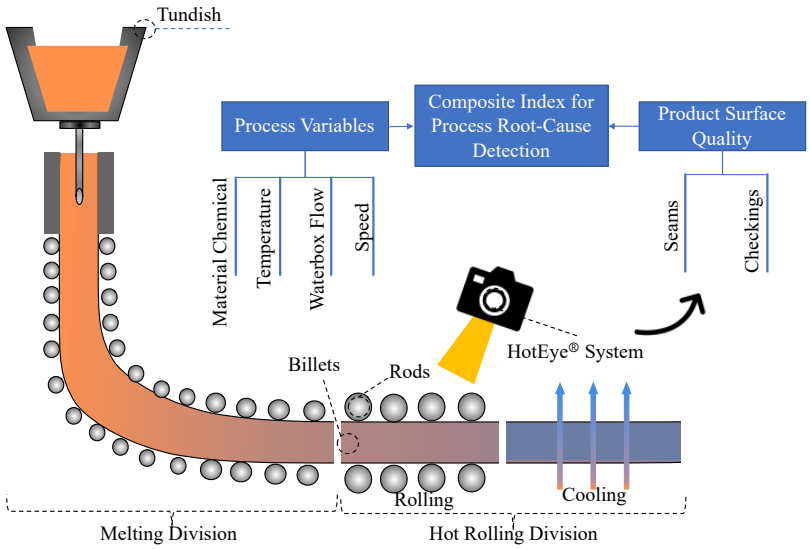
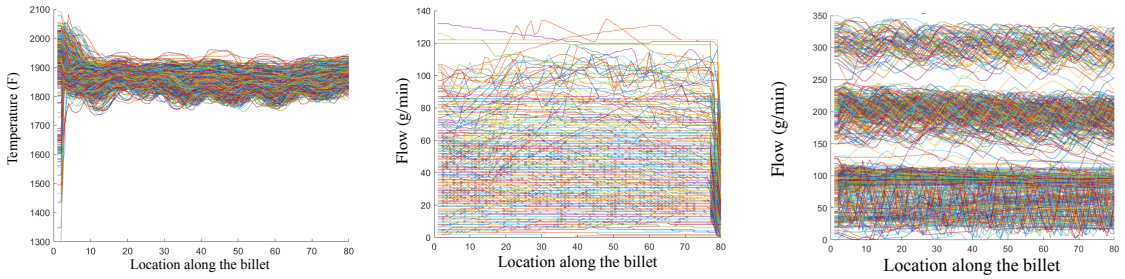


Figure 3.1 The Overview of Hot Rolling Process and the Data Generated from the System

non-twisting mill stand. They elongate the billet length and fine shape the billet surface, respectively. For a typical hot rolling process, there are around 10~20 intermediate stands, and 8~10 non-twisting mill stands installed. After the entire manufacturing process, a long thin steel billet is fabricated and coiled for transportation.



(a) Bottom temperature of the billets on stand 5

(b) Water flow on stand 10

(c) Water flow on station 26

Figure 3.2 Sample Measurements of the Process Variables

The recent development of the computer and sensing technology enables various types of process variables installed in the hot rolling process. During the hot rolling processes, process variables are collected from each station. These process variables are usually in the form of time-series data, representing the measurements taken from one end of the billet to the other end, as illustrated in Figure 3.2. Among them, the side and bottom temperature of the billet, the speed of cooling water shoot onto it (water flow), and its travel speed are collected when this billet is passing through an intermediate stand. Process variables are also collected as billet enters and exits the non-twisting mill stand. For example, the entry and exit speed, temperature and flow of cooling water applied to the billet are measured and collected before it travels into the non-twisting mill stand. After

the final hot rolling stand, an in-line optical sensor is equipped at the top of the production line, which measures the surface defects of the billets. Then, image processing software is used to count two types of defects, which are seams and checkings, on the billet's surface.

3.2.2 Challenges to the Conventional Approaches

The objective of data analysis in this chapter is to identify process variables collected from the hot rolling process that influence the quality variable of the product. This objective is achieved by analyzing the dependent relationship between the process variables and the product quality variables. Conventionally, the dependency between the process variables and product quality variable is described by models like logistic regression (Jin *et al.* 2007). However, these modeling methods only deal with a small number of process variables for a pre-specified form of dependency relationship. In the meantime, highly trained data scientists are needed to tune the model performance and interpret the model explicitly. Moreover, the modeling approach cannot identify all process variables that influence product quality if the process variables correlate with each other.

As illustrated in Section 3.2.1, a hot rolling process generates a wide variety of data with high complexity. From these data, we made the following specific observations. (1) There are multiple clusters in data sets generated from a hot rolling process. As an example, the water flow data acquired from station 26 (Figure 3.2 (c)) is naturally divided into three groups, due to multiple rolling conditions. (2) The signals are of distinct formats. Although the time series measurements from the process can be transformed to signals of the same length after preprocessing, these time series have distinct variability patterns, which can be seen from the temperature and the water flow rate measurements from two stations in

Figure 3.2. For such complex data, we typically need to transform the measurements from each variable into a set of features, which either capture the engineering knowledge of the processes or summarize the major variation patterns of the data. (3) In a hot rolling process, there are no systematic ways to identify the dependent relationships between the process variables and the product variables. As a result, a specific method may not be effective in identifying all possible dependency relationships between the process variables and product quality. In our manufacturing data analytics framework, this challenge will be addressed through the integration of multiple pre-specified methods. The pre-specified methods in our proposed framework are trained based on the standardized training procedures that can be done automatically by existing software or algorithms such as AutoML (Guyon *et al.*, 2015). Since our proposed method provides the influence status of the process variable according to nonparametric permutation tests, it does not need explicit knowledge to interpret the model and tune the parameters.

3.3 The Analysis Procedure of Manufacturing Data Engine (MADE)

In a hot rolling process, the raw data acquired from each process sensor is in the form of long time series. In the data preprocessing step, this long time series is cut into short ones that each corresponds to a rolling bar. Each short time series are aligned such that each of them has the same amount of measurements. Let there be N rolling bars obtained from the long signals. The measurements of all N products acquired from the process variable i are then organized as a data matrix $\boldsymbol{X}_i \in \mathbb{R}^{N\times J}$, $\mathrm{X_i} \in \mathrm{R^{n\times p_i}} i = 1, \dots, I$; where I is the total number of process variables. The n-th row of $\boldsymbol{X}_i$ is denoted by $\mathbf{x}_i^{(n)} = \left(x_{i1}^{(n)}, \dots, x_{iJ}^{(n)}\right) \in \mathbb{R}^J$ and it represents the short time-series measurements of

process variable i for product $n = 1, \dots, N$. The quality variable obtained from the metrology system for product n is denoted as $y^{(n)}$: $y^{(n)} = 1$ if the product n is defective, and $y^{(n)} = 0$ otherwise. The quality variable for all products are written as $\boldsymbol{y} = \left(y^{(1)}, \dots, y^{(n)}\right)$. These preprocessed measurements and quality variables are the starting point of the analysis procedure of MADE.

The analytical procedure of the MADE system has four major modules. First, MADE performs the clustering and outlier detection. Then, extract extensive features from the complex data of each stage. After that, evaluate the dependency measure between the features corresponding to each process variable and the product quality through multiple methods and nonparametric permutation tests. Finally, perform decision-level fusion by aggregating all dependency measures corresponding to every process variable and every method to identify the process variables that influence the final quality. The pre-specified input and output format of each module enable quick modification for specific processing steps according to the process requirements. We introduce each module in the subsections below.

3.3.1 Feature Extraction and Feature Selection

We first perform the feature extraction. A collection of k_i features are extracted from the measurements corresponding to process variable i. The feature k obtained from the process variable i for product n has a dimension of s_{ik}, and it can be written as $\mathbf{f}_{ik}^{(n)} \in \mathbb{R}^{1 \times s_{ik}}$. The collection of this feature for all products is given as $\mathbf{F}_{ik} = \left[\mathbf{f}_{ik}^{(1)\top} \cdots \mathbf{f}_{ik}^{(N)\top}\right]^{\top} \in \mathbb{R}^{N \times s_{ik}}$, $k = 1, \dots, k_i$. The approaches to extract these features $\mathbf{F}_{i1}, \dots, \mathbf{F}_{ik_i}$ are determined

during the system setup. In particular, these features fall into two categories, respectively selected by data analytical guidelines and engineering guidelines.

Data analytical guideline extracts the features based on the characteristics of the data. Below are some examples of the specified features corresponding to a process variable i.

- The mean of all measurements in the preprocessed time series gives a one-dimensional feature $\bar{x}_{i\cdot}^{(n)} = \left(\sum_{j=1}^{J} x_{ij}^{(n)}\right)/J$ for each product j and sample i;
- The sample variance of all measurements in the preprocessed time series gives another one-dimensional feature $s_i^{(n)} = \left[\sum_{j=1}^{J}\left(x_{ij}^{(n)} - \bar{x}_{i\cdot}^{(n)}\right)^2\right]/J$;
- The leading r_{pc} principal components (Jackson 2005) scores of $\boldsymbol{X}_i$ gives an r_{pc}-dimensional feature
- The energy values $\sqrt{a_1^2 + b_1^2}, \ldots, \sqrt{a_{r_{\mathrm{f}}}^2 + b_{r_{\mathrm{f}}}^2}$ calculated from the Fourier coefficients $a_1, \ldots, a_{r_{\mathrm{f}}}$ and $b_1, \ldots, b_{r_{\mathrm{f}}}$ corresponding to the leading r_{f} sine and cosine components of the curve $\left(x_{i1}^{(n)}, \ldots, x_{1J}^{(n)}\right)$ gives an r_{f} dimensional feature;
- The derivative of the curve $\left(x_{i2}^{(n)} - x_{i1}^{(n)}, \ldots, x_{iJ}^{(n)} - x_{i(J-1)}^{(n)}\right)$ gives an r_{d} dimensional feature.

Engineering guideline extracts the features specified by the engineering knowledge. In a hot rolling process, the surface quality problem of the products may be caused by uneven cooling of the billet surface, as indicated by the local variability of the

temperature and water flow. Therefore, the total variation and maximum derivative of curves are specified as two one-dimensional engineering-driven features.

- The total variation of a curve $\left(x_{i1}^{(n)}, \dots, x_{iJ}^{(n)}\right)$ is defined as $\sum_{j=1}^{J-1}\left|x_{i(j+1)}^{(n)} - x_{ij}^{(n)}\right|$;
- The maximum derivative is defined as $\max_{1\le r\le p-1}\left|x_{i(j+1)}^{(n)} - x_{ij}^{(n)}\right|$.

As the value of r_{pc}, r_{f} and r_{d} are specified, the feature extraction step automatically generates a set of features from these two guidelines for each process variable i.

3.3.2 Clustering and Outlier Detection

In this phase, the extracted features $\mathbf{F}_{i1}, \dots, \mathbf{F}_{ik_i}$ for all products are clustered and the outliers in them are identified. Jacques and Preda (2013) introduced several existing methods developed for functional data clustering and outlier detection. One of the most popular methods is distance-based outlier detection using clustering algorithms based on specific distances for functional data. Hierarchical clustering builds a hierarchy of clusters according to the dissimilarity between sets of observations (Rokach and Maimon 2005). According to the data characteristics, it is difficult to define a unified dissimilarity measurement for complex sets of data. In this chapter, the dissimilarity measurements used are the distance between sample means and the distance between sample derivatives.

In this chapter, the agglomerative hierarchical clustering method has been applied to the features that typically demonstrate clustering patterns. The agglomerative hierarchical clustering procedure corresponding to a feature $\mathbf{F}_{ik} \in \mathbb{R}^{N\times s_{ik}}$ is given in

Algorithm 3.1 (Kaufman and Rousseeuw 2009). The distance between two clusters A and B of these N products, $d(A,B)$, is defined as the shortest Euclidian distance between a sample in A and another sample in B, that is, $d(A,B) = \min_{m\in A, n\in B} \left\| \mathbf{f}_{ik}^{(m)} - \mathbf{f}_{ik}^{(n)} \right\|_2$. As the MADE system automatically generates $\mathcal{S}_1, \ldots, \mathcal{S}_N$, the clustering results with respect to 1, 2, …, N clusters, the process engineer then selects n^*, an appropriate number of clusters based on the visualization of the features' distributions, and generates the final result of clustering $\mathcal{S}_{n^*}$. Note that in the final result of clustering $\mathcal{S}_{n^*}$, the outliers are specified by the clusters that contain a single product.

After the clustering procedure, the labels of the derived clusters are appended as a new column to the features $\mathbf{F}_{ik}$'s obtained from the data analysis and engineering guidelines. To simplify the notations, we keep using $\mathbf{F}_{ik}$ to denote the kth feature corresponding to process variable i thereafter. The collection of all k_i features obtained from process variable i is written as $\mathbf{F}_i = \left[\mathbf{F}_{i1} \cdots \mathbf{F}_{ik_i}\right] \in \mathbb{R}^{N\times\sum_{k=1}^{k_i} s_{ik}}$, and those features for one product n is written as $\mathbf{F}_i^{(n)} = \left[\mathbf{f}_{i1}^{(n)} \cdots \mathbf{f}_{ik_i}^{(n)}\right]$. Clearly, $\mathbf{F}_i = \begin{bmatrix} \mathbf{F}_i^{(1)} \\ \vdots \\ \mathbf{F}_i^{(n)} \end{bmatrix}$.

Algorithm 3.1: Agglomerative hierarchical clustering

Initiate $\mathcal{S}_N = \{\{1\}, \ldots, \{N\}\}$.

For $m = N, \ldots, 2$:

Calculate the distance $d(A_m, B_m)$ between each pair of clusters $A_m, B_m \in \mathcal{S}_m$.

Select $(A_m^*, B_m^*) = \underset{\substack{A_m, B_m \in S_m \\ A_m \neq B_m}}{\arg\min}\ d(A_m, B_m)$

Set $S_{m-1} = S_m \setminus \{A_m^*, B_m^*\} \cup \{A_m^* \cup B_m^*\}$.

End For

3.3.3 Dependent Measure Evaluation

Recall that our final objective is to identify the variables and manufacturing steps that have a significant influence on the product quality. To achieve this objective, one needs to evaluate the dependency between the process features derived from the measurements of every process variable and the final quality of the product.

In general, if there exists a dependent relationship between the product quality and the process features, it is possible to predict the product quality with a higher accuracy using the process variables than to guess the quality of the product randomly. Therefore, whether the process features are related to the product quality can be inferred by evaluating the predictive performance of specific models. However, two challenges are involved here. First, we typically do not know the type of predictive method that achieves higher predictive accuracy. Second, the possible intercorrelation between measurements obtained from different sensors will affect the identification of influential process variables (Gunst 2018). Our strategy is thus to examine the dependent relationship between the features from individual process variables and the product quality variable using multiple predictive methods, such as logistic regression, classification, regression tree. After that, we will aggregate their evaluation results. In this section, we introduce an evaluation scheme based

on nonparametric permutation tests that generate a composite index called "dePendency INdicator (PIN)" describing the dependent relationship between the features corresponding to each process variable and the product quality.

Dropping the subscript i that denotes the index of the process variable, let $\mathbf{F}$ be the set of all features obtained from a process variable of N products, and $\mathbf{F}^{(n)}$ be the features of product n. Given a method that predicts the quality variable based on the features of process variable i, a prediction function $\hat{g}(\cdot)$ is obtained from the training data $\mathbf{F}$ and $\boldsymbol{y}$. To increase the predictive stability of the resulted predictive function, we first apply the ROSE method (Menardi and Torelli 2014) on $\mathbf{F}$ and $\boldsymbol{y}$ to generate a modified set of features $\mathbf{F}^*$ and quality variable $\boldsymbol{y}^*$ with balanced defective and non-defective samples. Then, we obtain the predictive function $\hat{g}(\cdot)$ from the modified data $\mathbf{F}^*$ and $\boldsymbol{y}^*$. Based on function $\hat{g}(\cdot)$, the quality of every product n is predicted through its feature $\mathbf{F}^{(n)}$

$$\hat{y}^{(n)} = \hat{g}(\mathbf{F}^{(n)}). \tag{3.1}$$

Recall that our objective is to obtain the PIN for every process variable given a predictive method, based on a multiple permutation test procedure. The multiple premutation test compares the predictive performance of a quality measure using $\mathbf{F}$ and $\boldsymbol{y}$, to the predictive performance of this method when the process features and product quality are independent. In this procedure, the predictive performance of the function $\hat{g}(\cdot)$ is measured by Matthews Correlation Coefficient (MCC) (Matthews 1975), a widely used performance metric in bioinformatics. We applied MCC as the measure for predictive performance because it is an appropriate measure for binary quality variable and thus fits

our application. Also, it is appropriate for the situation in which the defective products and the non-defective products are imbalance in size (Powers 2011). Thus, it is valid for many manufacturing processes in which the number of defective products is much smaller than the number of non-defective ones.

After obtaining the prediction $\hat{\boldsymbol{y}} = \left(\hat{y}^{(1)}, \ldots, \hat{y}^{(N)}\right)$ from the predictive model $\hat{g}(\cdot)$, the value of MCC for a process variable is defined as

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \tag{3.2}$$

In this expression, the number of correctly predicted defective products is denoted by $TP = \sum_{n=1}^{N} 1_{\hat{y}^{(n)}=1} 1_{y^{(n)}=1}$, standing for "true positive cases". Similarly, the number of non-defective products that are predicted correctly is denoted as TN, the number of non-defective products that are misclassified as defective ones is denoted as FP, and the number of defective products that are misclassified as non-defective ones is denoted as FN. The MCC value ranges from -1 to +1, and a large MCC value indicates that the function $\hat{g}(\cdot)$ has a good predictive performance when it is targeted for the training data **F** and $\boldsymbol{y}$.

In the multiple permutation tests, we approximate the distribution of MCC of a given predictive method. The null hypothesis is no dependency between the process variable and the quality variable. Such approximation is achieved by pairing the process features and the product quality randomly according to all samples, calculating the MCC values according to these randomly paired process features, and finally constructing the empirical distribution of all MCC values. Specifically, in the t^{th} permutation step, we

obtain the randomly-permuted feature set $\mathbf{F}^{\{t\}} = \begin{bmatrix} \mathbf{F}^{\left(p_1^{\{t\}}\right)} \\ \vdots \\ \mathbf{F}^{\left(p_N^{\{t\}}\right)} \end{bmatrix}$, where $\left(p_1^{\{t\}}, \dots, p_N^{\{t\}}\right)$ is a random permutation of $1, \dots, N$. Using $\mathbf{F}^{\{t\}}$ and $\boldsymbol{y}$ as the training data, we then derive the predictive function $\hat{g}^{\{t\}}$, so that we can predict the permuted features, $\hat{y}^{\{t\}(n)} = \hat{g}^{\{t\}}\left(\mathbf{F}^{\left(p_n^{\{t\}}\right)}\right)$ for all sample $n = 1, \dots, N$. Then the corresponding MCC value $MCC^{\{t\}}$ is calculated based on the prediction $\left(\hat{y}^{\{t\}(1)}, \dots, \hat{y}^{\{t\}(n)}\right)$ and $\boldsymbol{y}$ using formula (2). The distribution of MCC under the null hypothesis is sketched by the empirical distribution of all MCC values based on permutations.

If the predictive method indeed have a good performance, the absolute value of the MCC calculated from the real data set should be significantly higher than the absolute values of most permuted MCC values, $\left|MCC^{\{1\}}\right|, \dots, \left|MCC^{\{T\}}\right|$. Otherwise, the MCC and the MCC values from the permuted data follow the same distribution. Therefore, we define the PIN value as the empirical percentile of the absolute value of MCC (obtained from $\mathbf{F}$ and $\boldsymbol{y}$ directly without permutation) among $\left(\left|MCC^{\{1\}}\right|, \dots, \left|MCC^{\{T\}}\right|\right)$. It follows uniform distribution under the null hypothesis. However, if $\mathbf{F}$ and $\boldsymbol{y}$ are dependent and the predictive method is effective, the predictive error calculated from the original data should be significantly smaller than the predictive error from the randomly-paired samples, and thus the PIN value should be close to 0. Therefore, the derived PIN characterizes the dependency of the process features $\mathbf{F}$ and the quality variable $\boldsymbol{y}$, as measured by a given predictive method. The proposed dependent measure evaluation algorithm is summarized in Algorithm 3.2.

Algorithm 3.2: Dependency indicator evaluation for a process variable and a pre-specified prediction method

Inputs: process features **F**, product quality $\boldsymbol{y}$, permutation times (T)

Result: Dependency indicator of process variable (PIN)

Based on the data **F** and $\boldsymbol{y}$, use the pre-specified method to train a predictive function $\hat{g}(\cdot)$:

Estimate $\hat{y}^{(n)} = \hat{g}\ (\mathbf{F}^{(n)})$.

Calculate the MCC value based on $\left(\hat{y}^{(1)}, \dots, \hat{y}^{(N)}\right)$ and $\boldsymbol{y}$ from Equation (3.2).

For $t = 1, \dots, T$ **do:**

Select $\left(p_1^{\{t\}}, \dots, p_N^{\{t\}}\right)$, a random permutation of $1, \dots n$.

Based on the permuted data $\mathbf{F}^{\{t\}} = \begin{bmatrix} \mathbf{F}^{\left(p_1^{\{t\}}\right)} \\ \vdots \\ \mathbf{F}^{\left(p_N^{\{t\}}\right)} \end{bmatrix}$ and $\boldsymbol{y}$, train a predictive function $\hat{g}^{\{t\}}$.

Make prediction $\hat{y}^{\{t\}(n)} = \hat{g}^{\{t\}}\left(\mathbf{F}^{\left(p_n^{\{t\}}\right)}\right)$.

Get $MCC^{\{t\}}$ from Equation based on $\hat{\boldsymbol{y}}^{\{t\}} = \left(\hat{y}^{\{t\}(1)}, \dots, \hat{y}^{\{t\}(N)}\right)$ and $\boldsymbol{y}$.

End For

Return *PIN,* the empirical quantile of $|MCC|$ within $\{|MCC^{\{1\}}|, \dots, |MCC^{\{T\}}|\}$.

3.3.4 Decision-level Fusion and Dependent Process Variable Selection

Algorithm 3.2 in the previous section introduced how to evaluate whether a quality variable is dependent on features from a process variable using one method. In practice, the dependency between the process features and the product quality may be only identified by specific methods, due to the corresponding structure of the data. For example, a nonlinear relationship cannot be identified by methods based on linear assumptions such as logistic regression. Therefore, to draw accurate conclusions on dependency between process variables and product quality, it is desirable to calculate the PINs based on multiple methods for every process variable. In this section, we achieve this goal by aggregating all PINs that represent how the features from every process variable affect the product quality using multiple predictive methods.

Suppose that there are H predictive methods in total. Denote the PIN value calculated from process variable i, based on method h by PIN_{ih}. Based on the PIN values corresponding to all $i = 1, \dots, I$ and $h = 1, \dots, H$, we identify the leading process variables that are related to the product quality. Under the assumption that only a small number of process variables are related to the product quality, we apply the Holm procedure (Holm 1979) for this purpose. The Holm procedure calculates the adjusted PIN values in the following way: first, the PIN values for a certain method h in $\{PIN_{ih}: i = 1, \dots, I; h = 1, \dots H\}$ are sorted with ascending order, by $PIN_{(1)ih} \leq \cdots \leq PIN_{(I)ih}$. The adjusted PIN value for PIN_{ih} is defined as

$$\widetilde{PIN}_{ih} = \min\{\max\{(\mathrm{I} - q + 1)PIN_{(q)h} \mid PIN_{(q)h} \leq PIN_{ih}, q = 1, \ldots, I\}, 1\}. \qquad (3.3)$$

If $\widetilde{PIN}_{ih} < \alpha$ for some process variable i and method h, the i^{th} process variable is regarded as dependent with the quality variable, evidenced by method h. According to the Holm procedure, this criterion controls the probability of falsely identifying dependency between the process features and the product quality for any pair of process variable and method below α. Compared with the naïve method of selecting all I PIN values below some given thresholds, the application of the Holm procedure has a higher detection power, if I is large and only a small number of PIN values are not uniformly distributed.

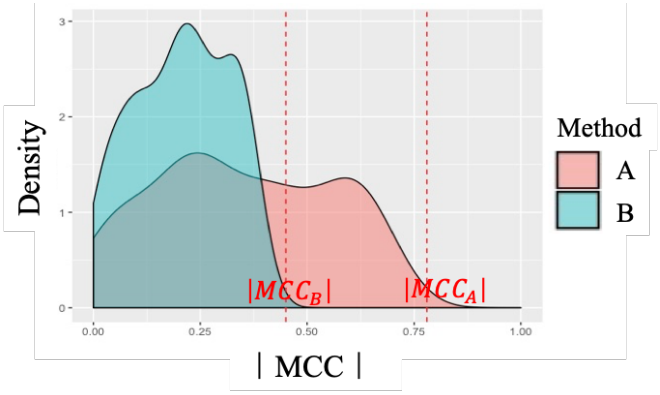


Figure 3.3 An Example of Comparable $\widetilde{PIN}_{ih}$ Value And Different MCC from Two Methods

The value of $\widetilde{PIN}_{ih}$ indicates whether the process variable i is dependent on the quality variable, according to the predictive method h. However, even if $\widetilde{PIN}_{ih}$ is small, the predictive performance of the method h on process variable i may still be poor, as indicated by the corresponding MCC value. As shown in Figure 3.3, both the value of $\widetilde{PIN}_{iA}$ and $\widetilde{PIN}_{iB}$ shows that process variable i is significantly dependent on the quality variable. However, the predictive performance of method A on process variable i (MCC_{iA})

is much lower than that of method B. We regard the process variable as being *influential* to the product quality, if it is related to the product quality, and it can predict the quality with good predictive performance measured by the MCC value. Therefore, we will identify the influential process variable by aggregating both measures of PIN and MCC to a composite index. We adopt the non-negative principal component analysis (NPCA) for this purpose, which has been used for composite index development for physician assessment (Liu *et al.* 2012).

The reason that we adopt NPCA instead of PCA is that the NPCA method constructs an interpretable composite index from PIN and MCC, as it requires that the loadings of the leading component be non-negative. This advantage is demonstrated with an illustrative example in Figure 3.4. In this example, the $\widetilde{PIN}$ and MCC of three methods are calculated for ranking the level of influences for the process variable on the quality variable. Method A is considered as a noneffective method for ranking the process variable dependency as it has the lowest MCC and a high $\widetilde{PIN}$ value. Method B provides evidence of rejecting the null hypothesis, as it has the highest MCC value and a small $\widetilde{PIN}$ value. Method C has a comparable small $\widetilde{PIN}$ value as Method B, but its MCC value is the worst

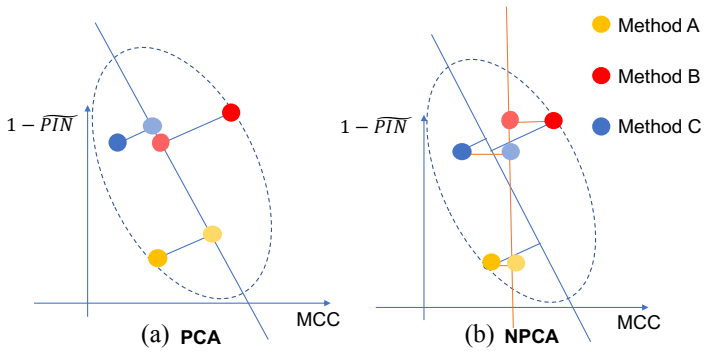


Figure 3.4 An Illustrative Example to Show the Necessity of using NPCA Approach for Dependency Ranking

among the three methods. Therefore, the ranking of these three methods should be $B > C > A$. As shown in Figure 3.4 (a), the first principal component of the MCC and PIN values leads to a ranking result of $C > B > A$, which is contradictory to the expectation. With the NPCA, by projecting the data points to the first non-negative principal component, the result of ranking is $B > C > A$, as shown in Figure 3.4 (b), which confirms our expectation. Therefore, we use NPCA to construct a composite index from the $\widetilde{PIN}$ and MCC to indicate the process variables that influence the product quality.

We slightly modify the formulations of NPCA to fit our problem. Recall that there are H methods and I process variables in total, and each pair of process variable i and method h corresponds to the PIN value $\widetilde{PIN}_{ih}$ and the MCC value MCC_{ih}. To perform the NPCA, we arrange them in a matrix $\mathbf{Z} = [\mathbf{z}_1\ \mathbf{z}_2] \in \mathbb{R}^{IH\times 2}$, where $\mathbf{z}_1 = \left(1 - \widetilde{PIN}_{11}, 1 - \widetilde{PIN}_{12}, \dots, 1 - \widetilde{PIN}_{IH}\right)^\top$ and $\mathbf{z}_2 = (MCC_{11}, MCC_{12}, \dots, MCC_{IH})^\top$**.** To find $\boldsymbol{\beta}_1$, which is the loading of the first non-negative principal component of $\mathbf{Z}$, we solve the following optimization problem:

$$\max_{\boldsymbol{\beta_1}} \boldsymbol{\beta}_1^\top \mathbf{Z}^\top \mathbf{Z} \boldsymbol{\beta}_1$$

$$s.t.\ \boldsymbol{\beta}_1^\top \boldsymbol{\beta}_1 = 1, \tag{3.4}$$

$$\boldsymbol{\beta}_1 \geq \mathbf{0}.$$

Denote the optimal solution of the problem (3.4) as $\boldsymbol{\beta}_1^*$**,** we obtain the corresponding PC scores as $\boldsymbol{S} = \mathbf{Z}\boldsymbol{\beta}_1^*$, whose l-th element is represented by S_l , $l = 1, \dots, IH$. Based on $\boldsymbol{S}$, the aggregated PIN for process variable i is calculated as $\ddot{\widetilde{PIN}}_i = \max_{l=1+(i-1)H,\dots,iH} S_l$, $i = 1, \dots, I$

which is the maximum value of the elements in $\boldsymbol{S}$ related to the process variable i. Large $\ddot{\tilde{PIN}}_i$ value indicates that the process variable i influences the product quality. After this procedure, the selected method h helps to scrutinize how the process variable relates to product quality. The above aggregation procedure is applied for ranking the process variables based on how much they influence the quality variable.

3.3.5 Summary and Implementation Guidelines

In summary, the analysis procedure of the MADE is illustrated in Figure 3.5. Note that in the entire analysis procedure, the practitioners can customize the analytical framework by replacing the clustering and outlier detection procedure, the feature extraction procedure, and the predictive methods, though replacing appropriate programming modules in the MADE system based on specific data format and the sensor layout of a system. The configurations introduced in the above sections are also specifically designed for analyzing the data from a hot rolling process. With customized configurations, the MADE system can be easily extended and adapted to new manufacturing or sensing systems.

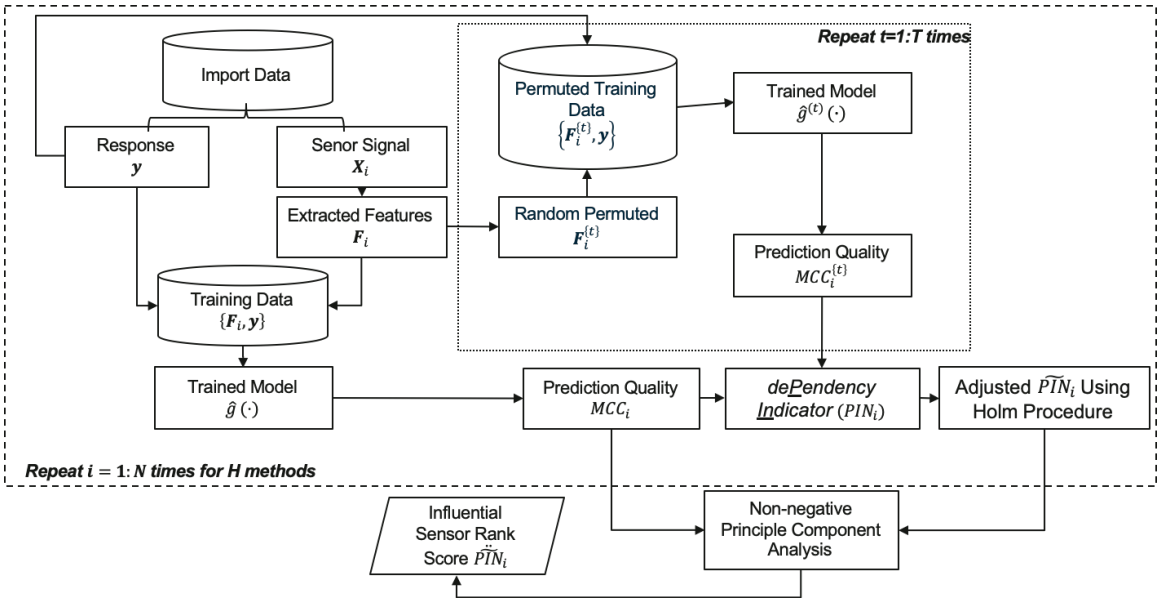


Figure 3.5 The Flowchart of the MADE System for a Hot Rolling Process

3.4 Numerical Analysis

3.4.1 Simulation Study

In this section, the performance of the MADE is evaluated through simulation studies. The objective of this simulation is to identify influential variables x_i $(i = 1,2,3,4)$ corresponding to the response y. In order to simulate the response dataset with influential variables and irrelative variables, we use "make_classification()" algorithm (Pedregosa *et al.* 2011) to generate two-class imbalanced response y and the corresponding influential variables (x_1 and x_2) and irrelative variables (x_3 and x_4). In particular, we have 20% of y are equal to one, 80% of y are equal to zero with 100 samples. The two influential variables are linearly correlated, i.e. $x_2 = 2x_1 - 1 + \epsilon$.

After we generate the dataset, we run the MADE for influential variable identification. In this simulation study, we jump over the feature extraction step by setting the extracted features F_i $(i = 1, 2, 3, 4)$ the same with variables x_i $(i = 1,2,3,4)$. Three tentative models, which are logistic regression, decision tree, and Support Vector Machine (SVM), are used to estimate the dependency between variables and response in 10,000 times permutation. Figure 3.6 shows the variables x_1 and x_2 are influential with the response y, which are consistent with our simulated facts.

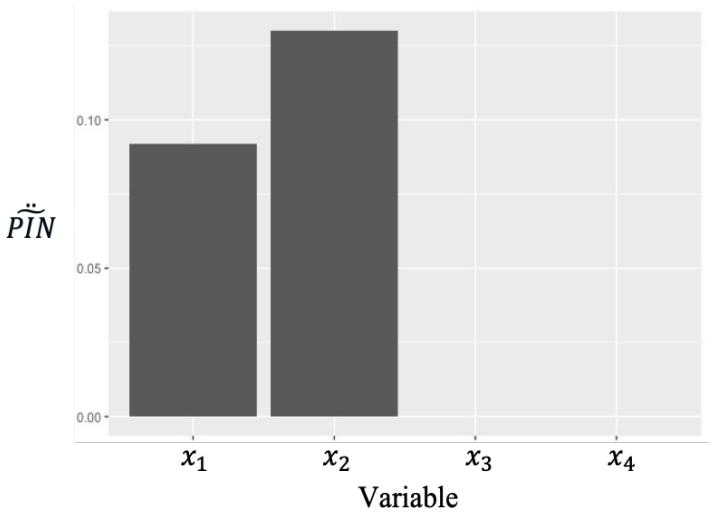


Figure 3.6 Influential Variables Identified by the MADE

3.4.2 Case Study

In this case study, we consider a hot rolling process, where an in-line optical sensing station, called HotEye® system (Chang *et al.* 2009a), is installed after the final rolling station. The HotEye® system provides real-time measurement of the surface quality of rolling bars by reporting various surface defects. In this case study, we focus on the seams on rolling bars, and our objective is to identify major process variables that influence the

occurrence of seams automatically. This is achieved by using the MADE system, whose detailed configuration and the corresponding results are introduced as follows.

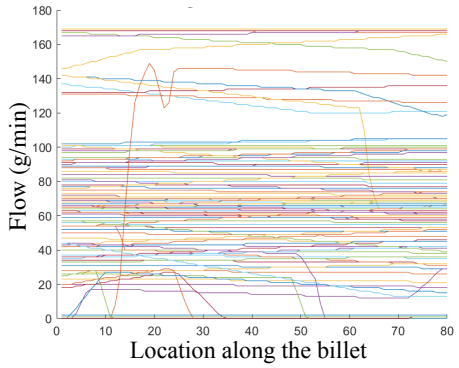
The data sets used in this case study are obtained from a hot rolling process. It includes the process data and quality measurements of 754 rolling bars. For each rolling bar, the in-situ process signals are transformed into a time series of length 80, and the preprocessed measurements from each process variable i form a 754×80 data matrix $\mathbf{X}_i$. The quality of each rolling bar is a binary variable, denoting whether each final product contains seams or not. The quality variables for all rolling bars are represented by vector $\boldsymbol{y}$ of length 754, containing zeros and ones.

3.4.2.1 Clustering, Outlier Detection and Feature Extraction

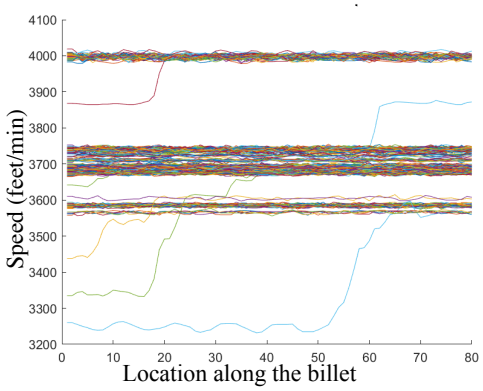
The original dataset is high dimensional functional curves. We first extracted the representative features from every process variable according to the procedure detailed in Section 3.3.1. As shown in Figure 3.7, the clustering of the process variables typically presents in the measurements of process variables 6, 11, 13 and 18. We, therefore, implemented the hierarchical clustering method for these process variables and added the derived group labels to each set of features. For example, the hierarchal clustering of the measurements for process variable 6 is based on the feature of “derivatives”, and the clustering of the measurements for process variable 18 is based on the feature of “mean value”. The features extracted for each process variable are listed in Table 3.1.

Table 3.1 Rolling Stage and the Corresponding Extracted Features

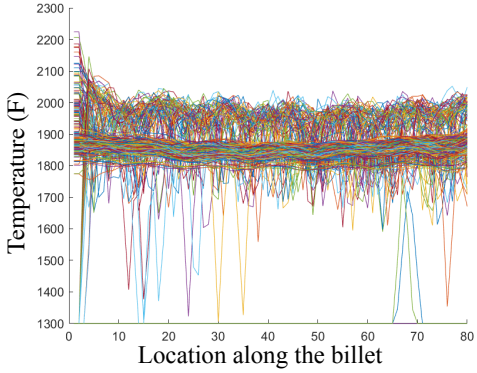
Index	Process variable name	Representative feature
3	stand5sidetemp	Mean, Energy, Variance
4	stand5bottomtemp	Mean, Energy, Variance
5	stand10flow	PCA score1, PCA score2, Mean, Total variation, Curvature
6	stand16flow	Group label, Mode, Total variation, Curvature
7	stand21temp	Mode, Total variation, Range
8	stand21speed	Mode, Total variation, Range
9	pntm1waterboxflow	Mean, Total variation, Curvature
10	pntm2waterboxflow	Mean, Total variation, Curvature
11	Ntmentryspeed	Group label, Mode, Total variation, Range
12	Ntmentrytemp	Mean, energy, variance
13	Ntmexittemp	Group label, Mean, Total variation, Curvature
14	stand26speed	Group label, Mean, Total variation, Curvature
15	s26waterbox1	Total variance, Range, Curvature, Median, Slope, Fitting error of Huber regression model
16	s26waterbox2	Mean, Total variation, Curvature
17	s26waterbox3	Mean, Total variation, Curvature
18	s26waterbox4	Group label, Mean, Total variation, Curvature.



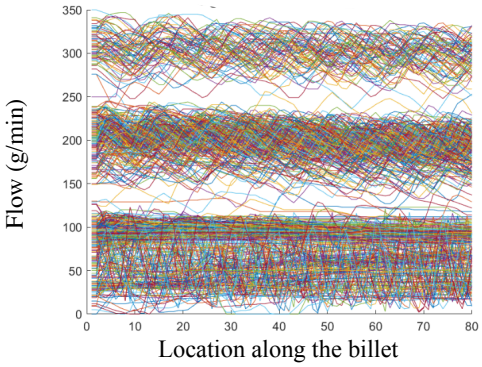
(a) Process variable 6



(b) Process variable 11



(c) Process variable 13



(d) Process variable 18

Figure 3.7 Clustered Process Variable Visualization

3.4.2.2 Dependent Measure Evaluation and Process Variable Selection

After obtaining the features for process variables, we implement the procedure discussed in Section 3.3.3 to calculate the PIN that represents the dependency between the measurements of process variable i and the product quality identified by three different predictive methods: logistic regression, support vector machine, and classification tree. The permutation times (T) is set to 10000. The aggregated $\widetilde{\widetilde{PIN}}$s for the seams are shown

in Figure 3.8. The decision-level fusion scheme introduced in Section 3.3.4 is used to calculate $\widetilde{PIN}$ and $\ddot{\widetilde{PIN}}_i$ as listed in Table 3.2.

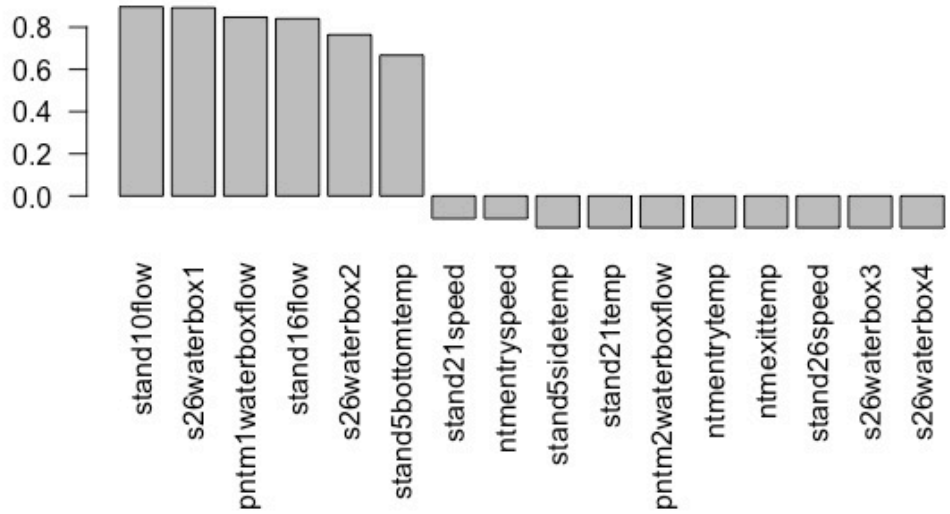


Figure 3.8 The Influential Stages Lead to Seams from the $\ddot{\widetilde{PIN}}$

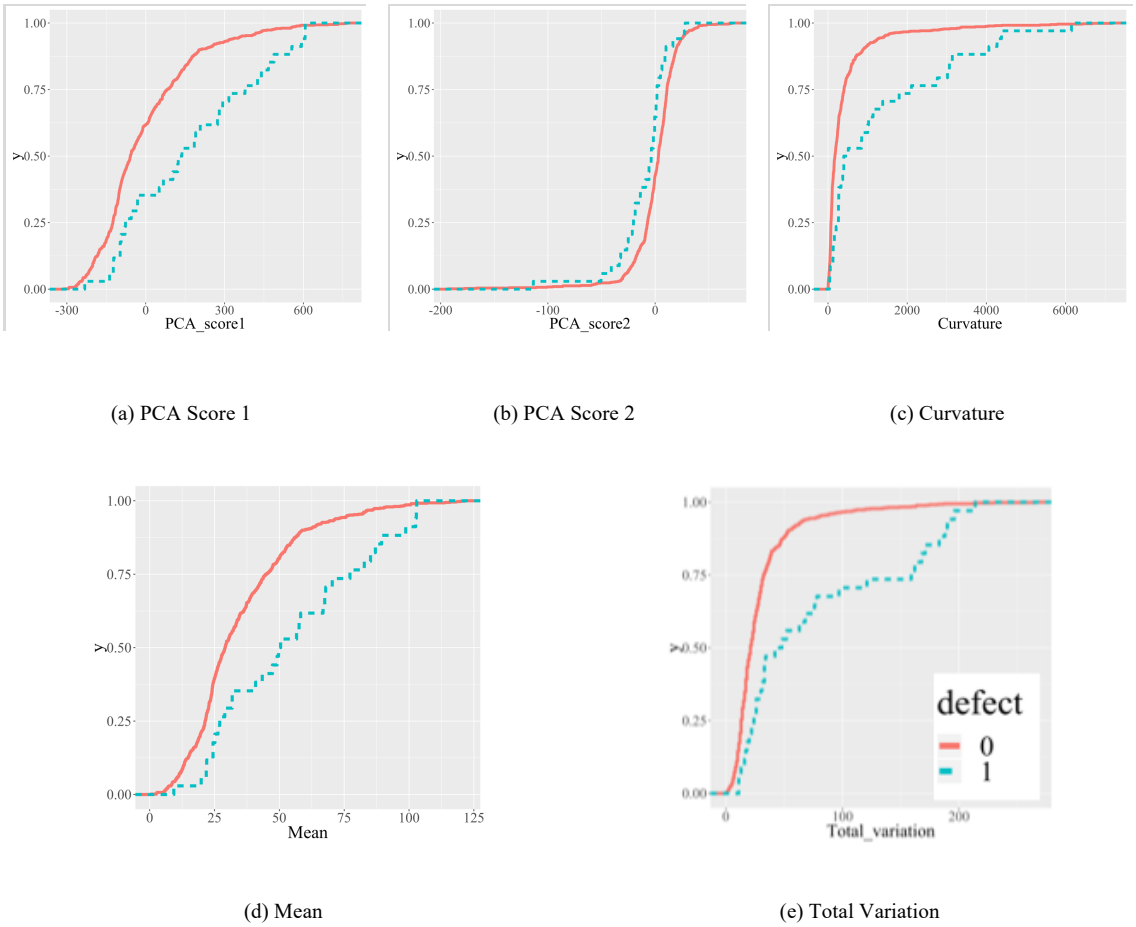
From the results shown in Figure 3.8, we see that the most influential process variables for seams on the product surface are "stand10flow" and "s26waterbox1".

3.4.2.3 Investigation of the Influential Level between the Process Variables and Product Defects

In this section, we validate that the selected process variable measurements and the corresponding features that derived from the process variables are indeed related to the product quality. We validate this by comparing the empirical distribution of the features for defective products that contain seams and the empirical distribution for non-defective products that do not contain seams.

Table 3.2 The Results of $\ddot{\widetilde{PIN}}$ for Seams

Process variable ID	Process variable name	$\widetilde{PIN}$			MCC			$\ddot{\widetilde{PIN}}$
		Logit Regression	Decision Tree	SVM	Logit Regression	Decision Tree	SVM	Seams
3	stand5sidetemp	1	1	1	0	0	0	-0.15
4	stand5bottomtemp	0.16	1	1	0.04	0	0	0.67
5	**stand10flow**	**0.51**	**1**	**0.02**	**0.14**	**0**	**0.39**	**0.9**
6	stand16flow	0.03	1	0.22	0.23	0	0.24	0.84
7	stand21temp	1	1	1	0	0	0	-0.15
8	stand21speed	1	1	1	0	0	0.17	-0.11
9	pntm1waterboxflow	1	1	0.03	0	0	0.24	0.85
10	pntm2waterboxflow	1	1	1	0	0	0	-0.15
11	Ntmentryspeed	1	1	1	0	0	0.17	-0.11
12	Ntmentrytemp	1	1	1	0	0	0	-0.15
13	Ntmexittemp	1	1	1	0	0	0	-0.15
14	stand26speed	1	1	1	0	0	0	-0.15
15	**s26waterbox1**	**1**	**1**	**0.01**	**0.03**	**0**	**0.34**	**0.89**
16	s26waterbox2	1	1	0.12	0.03	0	0.24	0.76
17	s26waterbox3	1	1	1	0	0	0	-0.15
18	s26waterbox4	1	1	1	0	0	0	-0.15



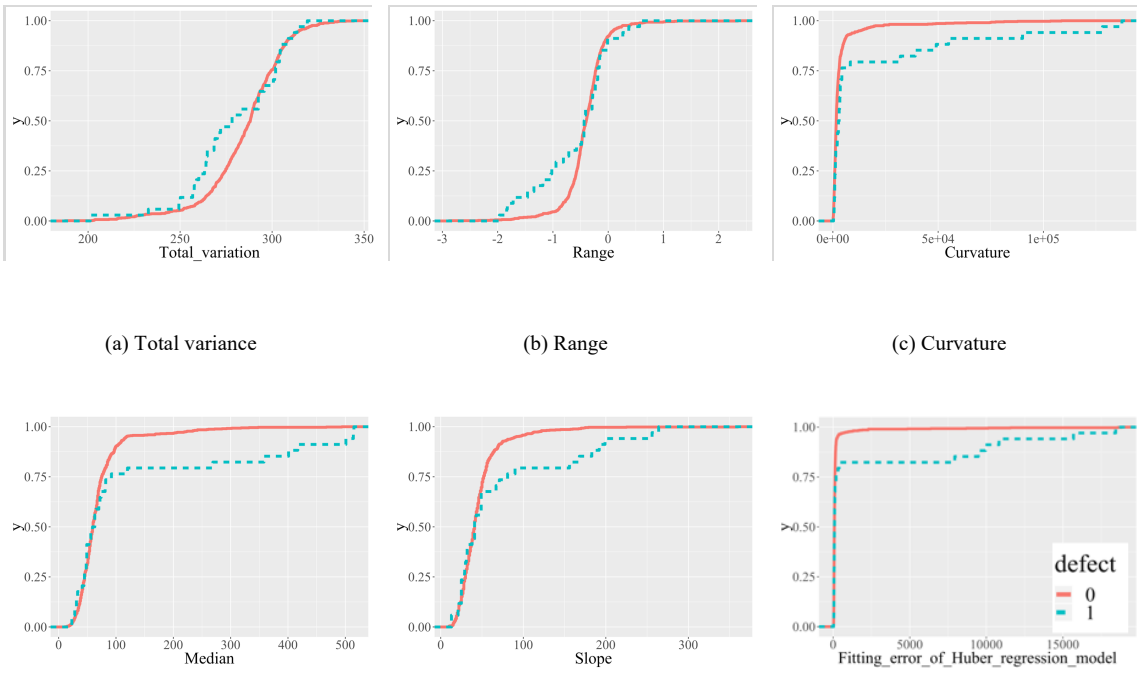
(a) PCA Score 1 (b) PCA Score 2 (c) Curvature

(d) Mean (e) Total Variation

Figure 3.9 Cumulative Distribution Functions (CDF) of Certain Features, for Defective (red) and Non-defective (blue) Products Obtained from Process Variable 5 (stand10flow)

The process variable 5 is the water flow at stand 10. Five features are derived from it, including the first PCA score, the second PCA score, curvature, mean and total variation. The empirical CDF of these individual features for defective products (with seams) and non-defective products (without seams) are shown in Figure 3.9. For this process variable, we find that the first PCA score (Figure 3.9 (a)) for the non-defective products is significantly smaller than that for the defective products. The second PCA score (Figure 3.9 (b)) for the non-defective products is greater than that for the defective products. The

curvature, mean, and total variation (Figure 3.9 (c, d, e)) of the non-defective products is smaller than that of the defective products.

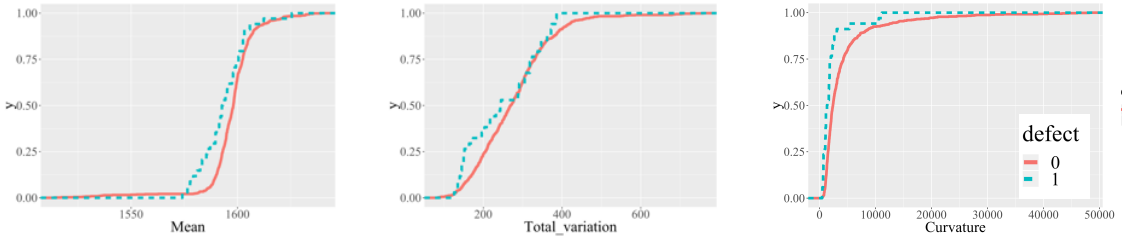


(a) Total variance (b) Range (c) Curvature

(d) Median (e) Fitting error of Huber regression model (f) Slope

Figure 3.10. Cumulative Distribution Functions (CDF) of Certain Features, for Defective (red) and Non-defective (blue) Products Obtained from Process Variable 15 (s26waterbox1)

The second influential process variables for seams is process variable 15 (s26waterbox1). As shown in Figure 3.10, the empirical cumulative distribution functions (CDF) of each feature for defective (with seams) and non-defective (without seams) products. Similarly, the non-defective group distinct significantly from the defective group for all features from the process variable 15.



(a) Mean (b) Total Variation (c) Curvature

Figure 3.11. Cumulative Distribution Functions (CDF) of Certain Features, for Defective (red) and Non-defective (blue) Products Obtained from Process Variable 14 (stand26speed)

To gain further understanding of the process variables that are not identified as the ones that influence the defect of seams, we investigate process variable 14 as an example. The process variable 14 represents the rolling speed at station 26, and three features are derived from it. The empirical CDF of these features for defective products (with seams) and non-defective products (without seams) are shown in Figure 3.11. Although the difference between the empirical distributions of extracted features is slightly different for defective and non-defective products, the gap is much smaller than the ones shown in Figure 3.9 and Figure 3.10.

The above observations demonstrate that the proposed MADE successfully identified the process variables that influence the seams defects of the rolling bars.

3.5 Conclusion

The proposed method focuses on an automatic analytical framework that streamlines the identification process of the influential process variables that are related to the quality variable. Since there is no similar system has been established, we did not compare the MADE to existing benchmarks. Our contribution is to build an integrated framework to automatically conduct initial influential process variable identification for a set of potential quality problems for the hot rolling process. The proposed framework automatically performs feature extraction, clustering, and automatically identifies the process variables that are influential to the product quality using multiple types of statistical models via multiple permutation tests.

There are several limitations of the MADE framework that calls for future studies. First, the proposed method only investigates the dependent relationship between each process variable and product quality. Therefore, it cannot detect if the interaction between process variables influences product quality. A quick solution for this problem is to add the interaction of features from multiple process variables to the multiple permutation tests. Second, the multiple permutation tests in MADE are computationally expensive. Modern parallel and distributed computing architectures can be employed to overcome this challenge. We refer the reader to the references (Kleiner *et al.* 2014; Fox *et al.* 2014) for more information.

CHAPTER 4. OPTIMAL PROACTIVE MAINTENANCE SCHEDULING FOR CLUSTER TOOLS IN SEMICONDUCTOR MANUFACTURING SYSTEM

4.1 Introduction

The semiconductor industry is highly competitive in terms of quality, productivity, and cost. The cluster tool, known as the wafer fabrication facility, is an automated robotic manufacturing system containing multiple computer-controlled process modules. A cluster tool, as seen in Figure 4.1, processes multiple substrates via load lock chambers for loading substrates, has a transfer chamber for transporting the wafers between the working chambers, and has several process chambers for performing one or more processes. The cluster tool is often described as a small factory for its multiple functions, such as chemical vapor deposition, physical vapor deposition, and plasma etching. With the capital investments for cluster tools as high as billions of dollars, the equipment downtime would bring considerable loss of productivity and profit (Perkinson *et al.* 1994). As a result, effective and comprehensive maintenance strategies are essential for improving equipment reliability and profit in a semiconductor fab.

A widely adopted maintenance scheduling approach with cluster tools is preventive maintenance, which is to conduct maintenance on a periodically prefixed schedule (Ben-Daya *et al.* 2016). However, this type of maintenance approach is a time- or machine-run-based schedule, which tends to overlook the reliability of the cluster tools. The current prevailing practices for maintenance scheduling of cluster tools are hierarchical planning

and scheduling approaches. These approaches usually contain two significant levels. The top-level is planning, which is for planning the operation time window between consecutive maintenance tasks. Given this time window, the lower level (e.g., the scheduling level) will arrange the exact time for each maintenance task. The concept of the two-level hierarchical model was proposed in (Scarf 2003). As an extension, (Yao *et al.* 2004) applied this model structure for preventive maintenance in semiconductor manufacturing.

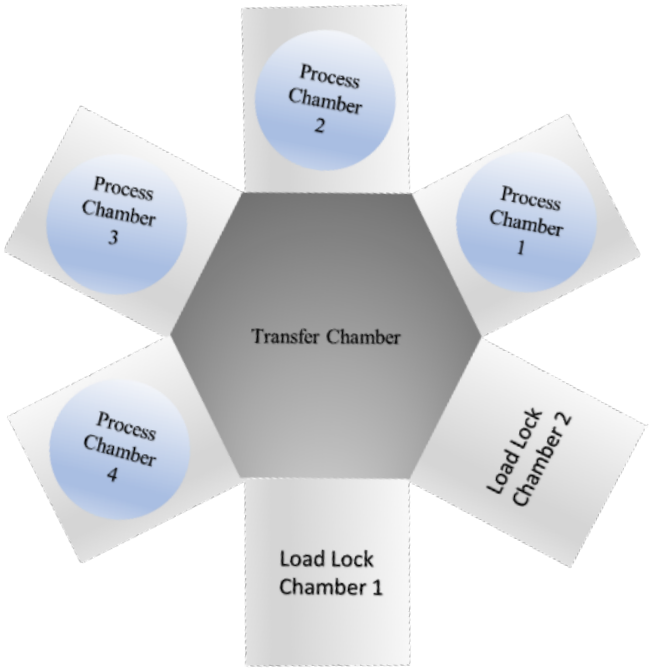


Figure 4.1 Cluster Tool Architecture

In this two-level hierarchical model, the top level of the two-level hierarchical modeling framework estimates an operation time window based on the deterioration status or the age of a machine. We noticed that the emergent maintenance operations might happen with a high probability after regular preventive maintenance operations. This situation could be caused by neglection of real-time health statuses of cluster tools. For

example, a mini fab has a cluster tool of chambers A and B, which have the same age and work in parallel, and A is much healthier than B. Under the two-level hierarchical modeling framework, chambers A and B will have equal priority for maintenance. If the maintenance task for A is before the task for B and meanwhile B is working for producing the product, then the failures may occur because of the degradation of B. The production line would, therefore, lose productivity during/after regular maintenance. Emergency maintenance operations would not only bring unexpected profit loss due to recourse shortage and labor inefficiency but also pay the "opportunity cost" (Eti *et al.* 2006). Hence, the degradation statuses of the chambers need to be considered to reflect the reliability of cluster tools.

To assess the reliability of the cluster tools, one needs to model the degradation statuses of the chambers in the cluster tools via in situ sensing data. The reason why we need to model the degradation statuses of the chambers rather than the cluster tool's is that the configurations of the chambers determine the reliability of the cluster tool. According to the recipes of products, the configurations of the chambers could be different over time. One needs to fuse degradation statuses according to the configurations for obtaining the cluster-tool-level reliability. The widely-adoption of sensing technologies in cluster tools and well-developed data analytic techniques remove the barriers of fusing sensing information to conduct in-situ assessment and estimation of system conditions. The conditions can be used to make maintenance decisions that help increase component operation life/availability and improve the profit rate for the plant.

On the other hand, machine availability is an essential indicator of an effective maintenance strategy. More specifically, the preventive maintenance decision making approaches mainly concerned about maximizing the availability of the cluster tools. The

availability of a cluster tool is the ratio of the capable being used time to the length of the total time interval. This availability has been very useful in terms of reflecting the cluster tools' performance in long term. However, the performance of cluster tools is not only affected by the long-term availability, but also by the real-time health condition, which is informed by the detailed sensing information. Degradation of the chambers of the cluster tools can lead to a domino effect that engulfs the whole process. For example, the plant would need to order the parts emergently since the factory has no inventory for the parts needed for emergency maintenance. Some other problems, such as workforce shortages, and eventually, order undeliverable, would also occur for emergency maintenance. Hence, it is necessary to proactively inspect the degradation trend to prevent cluster tools from running into a failure situation.

The proactive maintenance (Fitch 2013) strategy is a maintenance philosophy that seamlessly integrates diagnostics and prognostics information for maintenance decision making based on sensing information. A proactive maintenance plan gives a manufacturer the capability to prolong the life of machinery and prevent the unexpected breakdown of production. Though the concept of a proactive maintenance strategy is appealing, there are no systematic methodologies to implement the concept in a semiconductor manufacturing system with cluster tools. Towards this end, with consideration of the complexity of maintenance scheduling in semiconductor fabrication systems, we propose a three-level optimal proactive maintenance scheduling approach to schedule maintenance for the cluster tools in semiconductor manufacturing proactively.

The three-level optimal proactive maintenance scheduling approach has three hierarchical steps. The first step will provide a rough operation time window via analysis

of the demand pattern and the failure dynamics of the cluster tools. The operation time window is a period of time designated in advance to ensure minimal disruption to the planning production. The operation time windows allow clients of the service to prepare for possible disruption or changes. The second level is the AvaIlability-Degradation (AID) scheme. It measures the performance of the cluster tools with both the availability and the degradation. The third level is a Mixed-Integer-Programming (MIP) model that maximizes the overall profit rate by considering availability and tool degradation status, resource restrictions, as well as long-term planned operation time window for maintenance scheduling.

The main contribution of this study is to develop a practical, proactive maintenance scheduling approach that integrates degradation signals of chambers, availability of cluster tools, and resource allocation with consideration of the operation window. The optimized maintenance scheduling solution prevents the emergency (or reactive) maintenance that is caused by degradation and provides a method for maintenance scheduling of semiconductor cluster tools.

The rest of the chapter is organized as follows. Section 4.2 provides a brief literature review. Section 4.3 presents the overall proactive maintenance strategy and the three-level modeling framework. Section 4.4 gives a detailed discussion of the proposed AID scheme and the MIP model. A numerical case study is presented in Section 4.5 to demonstrate the efficiency of the proposed method. Section 4.6 concludes the chapter.

4.2 Literature Review

This section provides a review of related works in the literature. Since the proposed AID scheme integrates the degradation statue of machines into maintenance decisions, we will first provide a brief review of degradation estimation and modeling, and then discuss maintenance strategies for semiconductor manufacturing processes.

4.2.1 Degradation Estimation

The development of degradation modeling has been widely studied by using statistical and stochastic models. Lu and Meeker (Lu and Meeker 2012) first proposed nonlinear random-coefficient models for time-to-failure distribution estimation. Gebraeel *et al.* (2005) proposed a Bayesian degradation modeling approach to obtain real-time updated residual-life distributions (RLDs). Recently, degradation modeling with a neural network model (Gebraeel and Lawley 2007) and functional data analysis (Zhou *et al.* 2011) provided more opportunities to estimate the degradation status of complex systems such as semiconductor manufacturing.

A detailed review (Jardine *et al.* 2006) was provided on machinery diagnostics and prognostics with the consideration of sensing data. In Yildirim *et al.* (2016a), the authors proposed sensor-driven prognostic models that were integrated with sensor-driven RLDs. These updated RLDs captured the underlying state of degradation of system components using real-time sensor signals. Multiblock principal component analysis was used to extract features for the post-lithography process (Cherry *et al.* 2004), and modular artificial neural network combined with Dempster-Shafer theory was applied to detect faults in a plasma etching process (Hong *et al.* 2011). Given the multiple sensors deployed in a chamber, the multiple-profile sensor-based monitoring and anomaly detection scheme were developed

to detect the change point of the chamber (Zhang *et al.* 2018). The degradation trend can be estimated from the signal after the change point. Taking advantage of the development of degradation modeling methods, the proposed AID scheme has the flexibility to use different degradation modeling methods based on the plant sensors' characteristics. In this chapter, we mainly discuss how to integrate the chamber degradation status with the availability of the cluster tools rather than establishing the degradation models from the specific signals.

4.2.2 Maintenance Strategies for Semiconductor Manufacturing Processes

Several maintenance strategies for semiconductor manufacturing were investigated to establish further knowledge on the current maintenance issues. For a semiconductor fabrication process, a comprehensive maintenance schedule is required to account for the multiple cluster tools and the sensing data.

A common maintenance scheduling approach for semiconductor manufacturing is preventive maintenance, which is to conduct maintenance on a prefixed periodic schedule. Yao *et al.* (2004) proposed a calendar-based, two-layer hierarchical modeling framework on developing a decision-making strategy to maximize the profit rate with consideration of the availability and resource allocations. However, one limitation of Yao's approach is that it did not consider the degradation characteristics of individual components of the population. A Lagrangian decomposition coordination method was used to optimize the maintenance scheduling plan for a re-entering semiconductor facility (Kaihara *et al.* 2010). However, the problem setting under this method neglected the reliability of cluster tools as well. They assumed that the chamber of the cluster tools would not be down if it is not under maintenance. Hence, they believed that cluster tools can still work to produce products. However, cluster tools will break down due to the degradation of the chambers

even when they are not under maintenance. As a result, the cluster tools will suffer from an emergency breakdown due to the neglect of the chambers' degradation when applying this kind of maintenance strategies.

To address this issue, the 2015 International Technology Roadmap for Semiconductors (ITRS) highlighted the impact of predictive maintenance on the semiconductor sector. Conceptually, predictive maintenance includes fault detection and monitoring by performing statistical analysis on the sensing data, which is then used to determine the health conditions of the cluster tools. In this way, it avoids unnecessary replacements and optimizes inventory planning decisions. Machine learning methods were investigated to further improve predictive maintenance performance in semiconductor manufacturing (Luo *et al.* 2015; Susto *et al.* 2016.; Jalali *et al.* 2019; Butte *et al.* 2018). Moyne *et al.* (2016) presented various aspects of big data management technologies including data category, data storage, data treatment, feature selection, model building, model analysis, and cost-benefit analysis, and discussed how big data concepts could lead to potential solutions\tools for building maintenance strategies of semiconductor manufacturing processes.

The evolution from predictive maintenance to proactive maintenance for semiconductor manufacturing provides opportunities to improve overall equipment effectiveness and realizes the seamless integration of diagnostics and prognostics information with maintenance decision making, which is the objective of proactive maintenance (Fitch 2013). Given the opportunity by the development of computational and sensing technologies, several structural methods (Munirathinam and Ramadoss 2014; Muller *et al.* 2008; Bleakie and Djurdjanovic 2013) were built to demonstrate how to perform proactive maintenance for semiconductor equipment. However, those works lacked details beyond the generic modeling concepts.

In the next section, we will illustrate the proposed optimal proactive maintenance scheduling model, which considers the availability of cluster tools and the degradation of chambers and is applicable to real semiconductor fabricating processes.

4.3 Overview of the Optimal Proactive Maintenance Scheduling Model

In this section, we will present the optimal proactive maintenance scheduling model for semiconductor cluster tools. Generally speaking, the structure of the optimal proactive maintenance scheduling model includes three steps: (i) constructing a coarse operation time window from a planning model; (ii) incorporating the degradation of chambers with the availability of cluster tools by the sensor-driven AID scheme; and (iii) delivering the optimal maintenance schedule via the scheduling model.

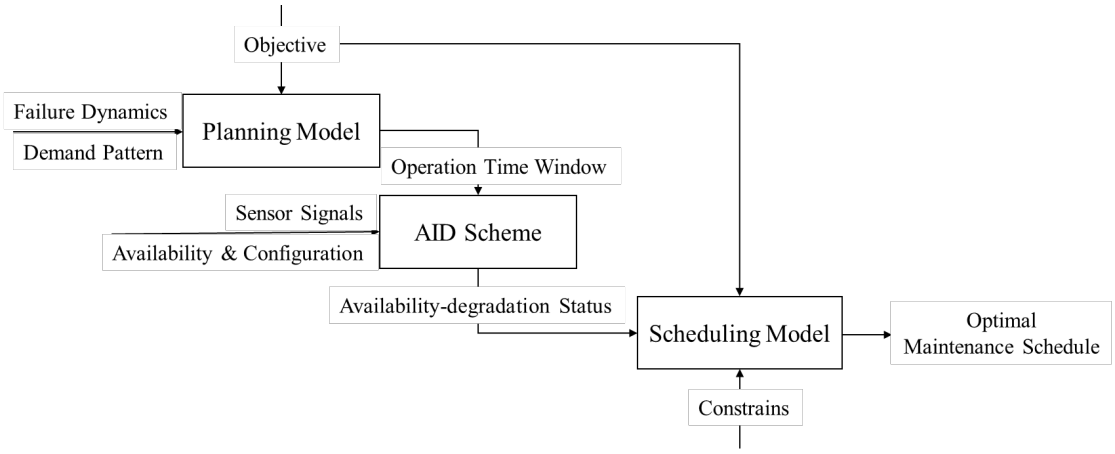


Figure 4.2 Optimal Proactive Maintenance Scheduling Model

4.3.1 Optimal Proactive Maintenance Scheduling Model

Our work is inspired by the widely adopted hierarchical planning and scheduling for semiconductor manufacturing maintenance. The proposed optimal proactive

maintenance scheduling approach has three components organized hierarchically. Figure 4.2 illustrates the structure of the three-level maintenance planning and scheduling approach.

- The first step is **planning model**. The purpose of the planning model is to obtain the optimal maintenance operation window for each maintenance task with respect to the failure dynamics and demand patterns for each cluster tool. A Markov Decision Process (MDP) (Puterman 2014) model is established to deal with the tool failure process and the incoming demand process for certain objective functions such as maximizing the total profit and maximizing the overall productivity. This method can be adopted to provide the operation time window for the next two steps. The operation time window obtained from the planning model acts like the coarse adjustment on the microscope, which provides a rough guideline for the Original Equipment Manufacturer (OEM) and the semiconductor fab to coordinate with each other to prepare for the maintenance. In this chapter, we mainly focus on the next two steps.
- The second step is the **sensor-driven AID scheme**. A critical component of our approach is an integration of the systematic AID scheme to access system performance for semiconductor manufacturing. With the maintenance tasks and their operation windows, one can estimate the AID status while doing each maintenance task, and the AID status provides references for the scheduling. Figure 4.3 presents the structure of the AID scheme. At the chamber level, we derive the degradation path by diagnosing the sensor-data. Then, we acquire the chamber health condition in terms of the surviving rate from the degradation model and then integrate it with the availability corresponding to different

maintenance tasks. The configurations of cluster tools vary based on the maintenance tasks undertaken, which include the required maintenance tasks and the corresponding operation time window. The plant-level AID value is synthesized with the cluster-tool-level AID value and the cluster-tool-level maintenance requirements. For the configurations of cluster tools, they will be introduced in Section 4.3.5. The details about how to obtain the AID will be discussed in the next section.

- **The scheduling model**, as the third step, carries out the optimal schedule to perform maintenance according to the given AID value, resource constraints,

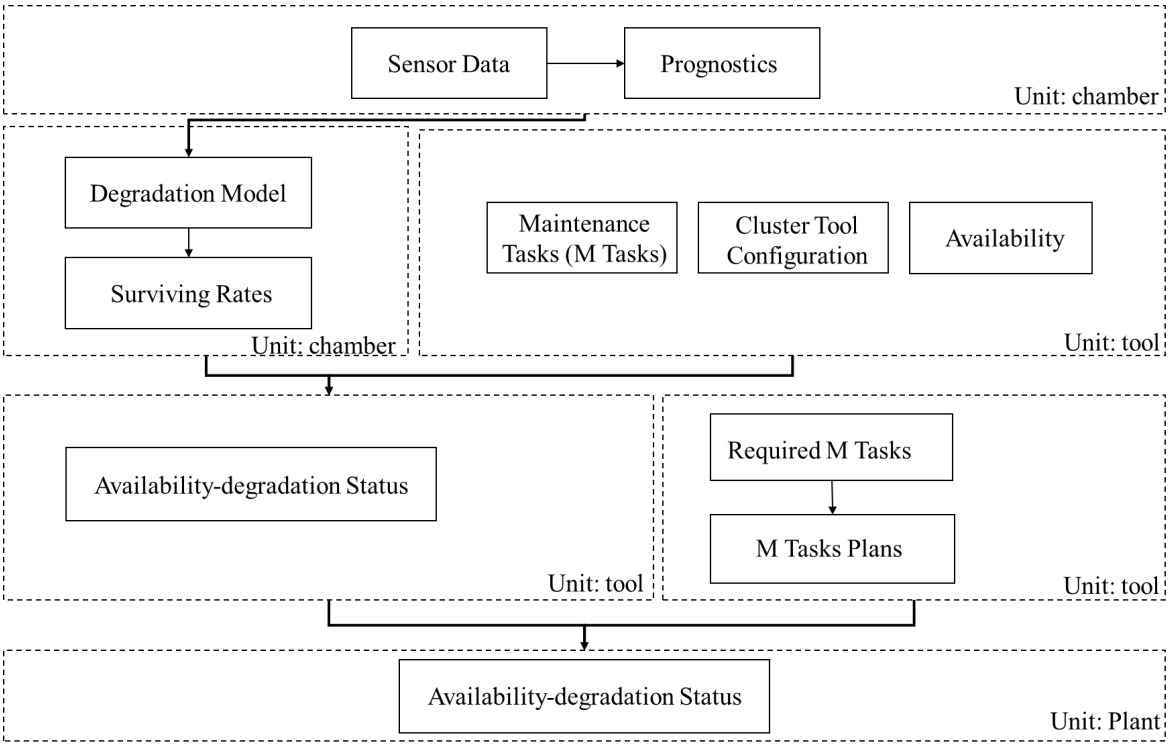
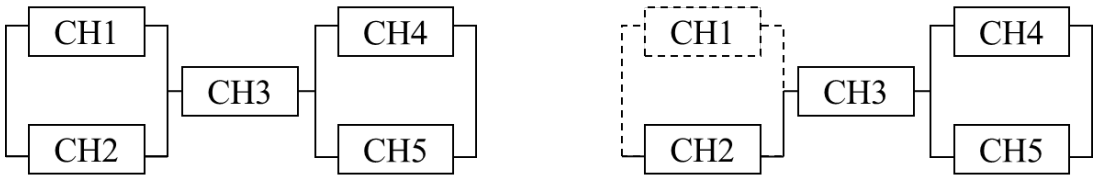


Figure 4.3 The Flowchart of the AID Scheme

and maintenance cost. The objective of this model is to maximize the profit rate during maintenance. A MIP model is established for the problem.



(a) The Physical Layout of a Cluster Tool (b) Configuration When Chamber 1 (CH1) is under Maintenance

Figure 4.4 Chamber Layout Examples of a Cluster Tool

4.3.2 System Description

The proposed optimal proactive maintenance scheduling model considers the degradation of a chamber that is monitored using multiple sensors. Under the same cluster tool configuration in Yao *et al.* (2004), Figure 4.4 (a) shows the physical layout of a cluster tool. We can see that the chambers (CH) within a cluster tool interact with each other. Hence, the behavior of each chamber depends on the behavior of others. When CH1 is under maintenance, the tool configuration will change to CH2 connected with CH3 in series, which is then connected with a parallel module of CH4 and CH5, as shown by Figure 4.4 (b). As each configuration represents a different tool level availability, we are now able to acquire the cluster-tool-level AID status.

Table 4.1 Configuration Matrix for a Cluster Tool (Legend: 0-down, 1-up, X-irrelevant) (Yao *et al.* 2004).

Scenarios	CH 1	CH 2	CH 3	CH 4	CH 5	Availability
1	1	1	1	1	1	100%
2	0	1	1	1	1	60%
3	1	0	1	1	1	60%
4	0	0	X	X	X	0%
5	1	1	1	0	1	80%
6	1	1	1	1	0	80%
7	X	X	X	0	0	0%
8	1	0	1	1	0	60%
9	1	0	1	0	1	60%
10	0	1	1	1	0	60%
11	0	1	1	0	1	60%
12	X	X	0	X	X	0%

Table 4.1 is a configuration matrix for a cluster tool provided in (Yao *et al.* 2004). The first row represents the scenario when all chambers (1–5) are up, indicated by "1," and cluster tool availability is 100%. The second and third rows represent the scenarios when either CH1 or CH2 is shut down, indicated with "0" for maintenance, respectively, with tool availability of only 60%. However, the fourth row shows that the tool availability is 0 when both CH1 and CH2 are down at the same time, regardless of the status of other chambers. This is because each wafer goes through either CH1 or CH2. When both chambers are down, no wafers can be processed. This also indicates that it is unwise to consolidate maintenance tasks for CH1 and CH2 because the tool availability will be zero. Similarly, the last row suggests that each wafer must go through CH3. When CH3 is down,

the corresponding cluster tool would have no throughput, and the cluster tool's availability is 0.

4.4 Development of the Optimal Proactive Maintenance Scheduling Model

The optimal proactive maintenance scheduling model is the AID integrated maintenance scheduling model that combines degradation signals from multiple sensors and maintenance scheduling decision-making into a MIP model. The extraction of degradation signals, AID parameters estimation, and the MIP model will be discussed in this section.

4.4.1 Data-level Degradation Modelling

Real-time sensing information is introduced to estimate the availability-degradation status of cluster tools. This status depends not only on chambers' configuration but also on the degradation status of the chambers of the cluster tools.

The profile data of q^{th} sensor, which measures the p^{th} sample at chamber c of i^{th} cluster tool, is treated as a continuous function $Y_{pq}^{i,c}(t)$ with respect to $t \in \mathcal{T} = [\tau_0, \tau_1]$, where $[\tau_0, \tau_1]$ is the production time interval. $Y_{pq}^{i,c}(t)$ can be formulated as

$$Y_{pq}^{i,c}(t) = \eta_q^{i,c}(t) + \epsilon_{pq}^{i,c}(t) \tag{4.1}$$

where $\eta_q^{i,c}(t)$ is the function of the template profile of the q^{th} sensor on the i^{th} cluster tool's chamber c and $\epsilon_{pq}^{i,c}(t)$ is the stochastic error with $E\left(\epsilon_{pq}^{i,c}(t)\right) = 0$ for $p = 1, \dots P;\ Q = 1, \dots, Q$. Then the profile data of all sensors can be denoted as $\boldsymbol{Y}_p^{i,c}(t) =$

$\left[Y_{p1}^{i,c}(t), \ldots, Y_{pQ}^{i,c}(t)\right]^T$, with $\boldsymbol{\eta}^{i,c}(t) = \left[\eta_1^{i,c}(t), \ldots, \eta_Q^{i,c}(t)\right]^T$, and $\boldsymbol{\epsilon}_p^{i,c}(t) = \left[\epsilon_{p1}^{i,c}(t), \ldots, \epsilon_{pQ}^{i,c}(t)\right]^T$. The subsequent i^{th} on-line testing sample after preprocessing, $\boldsymbol{Y}_p^{i,c}(t)$, is assumed to follow a change-point model

$$\boldsymbol{Y}_p^{i,c}(t) = \begin{cases} \boldsymbol{\eta}^{i,c}(t) + \boldsymbol{\epsilon}_i(t), & \text{for } i = 1, \ldots, \tau \\ \overline{\boldsymbol{\eta}}^{i,c}(t) + \boldsymbol{\epsilon}_i(t), & \text{for } i = \tau + 1, \ldots \end{cases} \quad \textbf{(4.2)}$$

where τ is the unknown change point, $\boldsymbol{\eta}^{i,c}(t)$ is the mean function before degradation, and $\overline{\boldsymbol{\eta}}^{i,c}(t)$ is the mean function along with degradation, which is also called a degradation path.

Since the sensor deployment of a complex manufacturing process is elaborate, the data has a relatively high dimension. Due to the cross-correlated property of the sensors, the correlation structure is analyzed to fuse the sensors with high correlations into one cluster. The detailed procedures of estimating $\overline{\boldsymbol{\eta}}^{i,c}(t) = \left[\bar{\eta}_1^{i,c}(t), \ldots, \bar{\eta}_Q^{i,c}(t)\right]^T$ can be found in (Zhang *et al.* 2018).

The surviving rate of the chamber is estimated from its degradation path. The specific method for estimating the degradation path varies case by case. We use a linear degradation case as an example to illustrate how to calculate the surviving rate of the chambers and clarify the notations. By assuming error terms from i.i.d. random error process, one could employ the following method. Let $\theta \sim N(\mu, \sigma^2)$ with $\sigma \ll \mu$ so that $P(\theta \leq 0)$ is negligible, ϕ is a known constant, and D is the failure threshold. Then for a linear degradation path of the i^{th} cluster tool's chamber c given by $\mathrm{y}^{i,c} = \bar{\eta}^{i,c}(t, \phi, \theta) + \varepsilon$, where $\bar{\eta}^{i,c}(t) = \phi + \theta t$ is the actual degradation path, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ is the measurement error, we can derive the surviving distribution of the i^{th} cluster tool's chamber c, $\bar{F}_i^c(t)$, as the following

$$\bar{F}_i^c(t) = 1 - \Phi\left(\frac{t-\frac{D-\phi}{\mu}}{\frac{t\sigma}{\mu}}\right) = 1 - \Phi(\mathrm{g(t)}) \quad \textbf{(4.3)}$$

where $\mathrm{g(t)} = \frac{t-\frac{D-\phi}{\mu}}{\frac{t\sigma}{\mu}}$. Residual life distribution of chamber c of the cluster tool i is $F_i^c(t) = 1 - \bar{F}_i^c(t)$.

4.4.2 AID Scheme Formulation

The degradation status of the chambers will significantly affect the reliability of the cluster tools while performing maintenance. Therefore, it is unneglectable for one to make maintenance decision without taking the degradation status into count. With the consideration of the degradation status of each chamber, the decision should not only rely on the availability of each configuration, as shown in Table 4.1, but also take the surviving rate of chambers that are not under maintenance into account. The AID value for cluster tool i is its availability with the probability that the chambers not under maintenance would not fail at that time. The AID value can also be regarded as the expected availability of the cluster tool. It is defined as

$$\mathrm{AID}_i(t) = V_i(t) \times \prod_i P(A) = V_i(t) \times \Psi_i\left(\bar{F}_i^1(t), \bar{F}_i^2(t), \dots, \bar{F}_i^{\mathcal{C}_i}(t)\right) \quad (4.4)$$

where $V_i(t)$ is the availability of cluster tool i at time t, A denotes the event that chamber i is not under maintenance and has not reached the failure threshold, $P(A)$ is the probability of event A occuring, $\Psi_i(\cdot)$ is the structure-function that depends on the configuration of the cluster tool i corresponding to maintenance tasks, $\bar{F}_i^c(t)$ is the surviving rate for the chamber c $(c = 1, 2, \dots, \mathcal{C}_i)$ of the cluster tool i. If two working

chambers ($c = 1, 2$) of the cluster tool i work in series, the structure-function is $\Psi_i\left(\overline{\boldsymbol{F}}_i(t)\right) = \bar{F}_i^1 \times \bar{F}_i^2$. If the two chambers work in parallel, the structure-function $\Psi_i\left(\overline{\boldsymbol{F}}_i(t)\right) = \left(1 - F_i^1(t)F_i^2(t)\right)$. For the case that the parallel chambers ($c = 1, 2$) in series with a chamber ($c = 3$), the structure-function would be $\Psi_i\left(\overline{\boldsymbol{F}}_i(t)\right) = \left(1 - F_i^1(t)F_i^2(t)\right) \times \bar{F}_i^3$. Based on this rule, we can formulate any configurations. For example, for the structure in Figure 4.4 (a), referenced from Table 4.1, we can obtain the structure-function as $\Psi_i\left(\overline{\boldsymbol{F}}_i(t)\right) = \left(1 - F_i^1(t)F_i^2(t)\right) \times \bar{F}_i^3 \times \left(1 - F_i^4(t)F_i^5(t)\right)$.

4.4.3 MIP Model Formulation

Let t denote a generic time-period, or maintenance decision epoch, and T the planning horizon ($t = 1, \dots, T$). M cluster tools are being considered in the model, and their indices are $i = 1, \dots, M$. For the i^{th} cluster tool, there would have ρ_i maintenance tasks with index $l = 1, \dots, \rho_i$. N recourses types, such as workforce and parts, are considered with index $j = 1, \dots, N$. $R^j(t)$ denotes the number of resource types available in period t for $j = 1, \dots, N$. $r_i^j(\cdot)$ is the resource function calculating the requirement of resource type j for the cluster tool i. The duration of the maintenance task l of the cluster tool i is d_i^l. The operation window associated with the maintenance task l of the cluster tool i is indicated by a binary variable $w_i^l(t)$. $w_i^l(t) = 1$ means that the maintenance task l of cluster tool i could be performed at time period t and $w_i^l(t) = 0$ otherwise, i.e.,

$$w_i^l(t) = \begin{cases} 0 & do\ not\ do\ maintenance \\ 1 & do\ maintenance \end{cases} \tag{4.5}$$

To perform maintenance task l on cluster tool i, the cost is c_i^l. To determine whether the maintenance task l on the cluster tool i in period t should be done, a binary decision variable is introduced to the model, i.e. $a_i^l(t)$. Therefore, the chambers' down/up status can be represented by $a_i(t)$, where $a_i(t) = \sum_{l_i} a_i^l(t)$, and l_i are the maintenance tasks on cluster tool i.

$$a_i^l(t) = \begin{cases} 0 & do\ not\ do\ maintenance \\ 1 & do\ maintenance \end{cases} \tag{4.6}$$

After the setting of indices, decision variables, and parameters, our model is shown in (4.7)-(4.15).

$$\max_{a_i^l} \sum_{t=1} \sum_{i=1} \left(b_i AID_i(t) - \sum_l c_i^l a_i^l(t) \right) \tag{4.7}$$

s.t.

$$\frac{\sum_{t=1}^{T} a_i^l(t)}{d_i^l} = 1 \tag{4.8}$$

$$a_i^l(t) \geq a_i^l(t-1) - \frac{1}{d_i^l} \sum_{s=t-d_i^l}^{t-1} a_i^l(s) \ \ for\ t > 1 \tag{4.9}$$

$$\frac{\sum_{t=1}^{T} a_i^l(t) w_i^l(t)}{d_i^l} - 1 \geq 0 \tag{4.10}$$

$$\mathrm{AID}_i(t) = V_i(t) \times \Psi_i \left(\bar{F}_i^1(t), \bar{F}_i^2(t), \dots, \bar{F}_i^{N_i}(t) \right) \tag{4.11}$$

$$V_i(t) = f_i \left(a_i(t), a_i(t-1), \dots, a_i\big(t-(k_i-1)\big) \right) \tag{4.12}$$

$$R^j(t) \geq \sum_{i=1}^{M} r_i^j \left(a_i(t), a_i(t-1), \dots, a_i\big(t-(k_i-1)\big) \right) \tag{4.13}$$

$$for\ i = 1, \dots, M, t = 1, \dots, T; a_i(t) = 0\ for\ t \leq 0$$

$$a_i^l(t) \in \{0,1\}, \forall t, i, l \tag{4.14}$$

$$w_i^l(t) \in \{0,1\}, \forall t, i, l \tag{4.15}$$

Achieving the factory's profit requires high cluster tools' availability and reliability. Therefore, the AID value is used as a performance measurement of the cluster tools. To maximize the total profits, the model uses the profits from the AID value of the cluster tools minus the costs of performing the operation tasks, which is equation (4.7). In equation (4.7), the profit rate b_i and maintenance cost c_i^l are parameters that need to be specified by the factory. The profit rate b_i is the profit per unit time when one cluster tool is operating with 100% availability. The maintenance cost c_i^l is the cost per unit time corresponding to the standby chambers under maintenance task l for cluster tool i. Equations (4.8-4.10) state that the scheduled operation tasks should be performed within their time windows. Equation (4.11) computes the AID value for each cluster tool for each period. Equation (4.12) computes the availability for each cluster tool for each time period. The sequence of $a_i(t), a_i(t-1), \dots, a_i\big(t-(k_i-1)\big)$ determines a row of the "configuration matrix", and the value of $f_i(\cdot)$ would be the respective availability value of the row. For the resource constraint, which is Equation (4.13), all cluster tools' resource demand over each type of resources must be less than the available type of resources in each time unit. The resource function $r_i^q(\cdot)$ depends only on the maintenance task vector $a_i(t), a_i(t-1), \dots, a_i\big(t-(k_i-1)\big)$.

4.5 Simulation Case Study

In the case study, we will implement the AID integrated proactive maintenance scheduling model and to evaluate the model performance. Following the simulation in (Yao *et al.* 2004), 11 cluster tools are working homogeneously with the same processing steps and configurations, i.e., M=11. The availability and configuration of these cluster tools are given in Table 4.1. Based on domain knowledge and clustering results in Zhang *et al.* (2018), we select S13 as an informative sensor for analysis. The extracted degradation signal is shown in Figure 4.5.

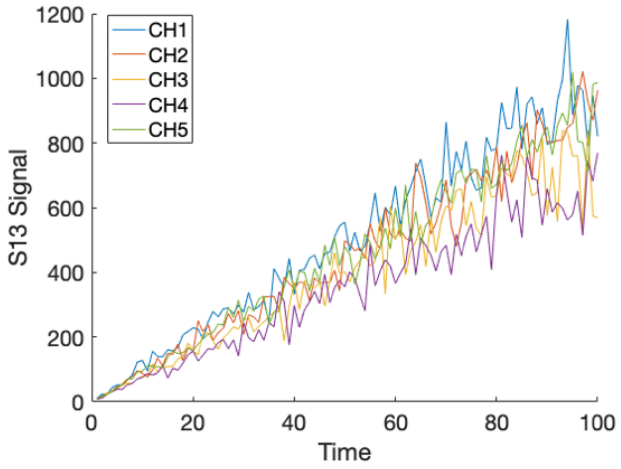


Figure 4.5 Degradation Signal Extracted from Chamber Degradation Signal

Based on the extracted degradation signal, we can obtain the AID value for one cluster tool under different maintenance tasks, which is shown in Figure 4.6, where the vertical lines represent the operation duration from the start (OS) to the end (OE) – 7 days in our case. There are 11 maintenance tasks in total. Their configuration matrix is given in Table 4.1. Since the AID value depends on the configuration of the cluster tool and the degradation of chambers, one could see that the AID values under the different tasks within the operation time window are distinct from each other. If the scheduling approach neglects the degradation status but only considers the availability, task 1 and task 3 would have the

same availability value and the same maintenance priority. However, from Figure 4.6, we could find out that the AID value corresponding to task 1 is smaller than the AID value corresponding to task 3. In this way, task 1 should have a higher priority than task 3. The structure-function of maintenance task vectors is listed as a library of cluster tool configurations that correspond to the tasks.

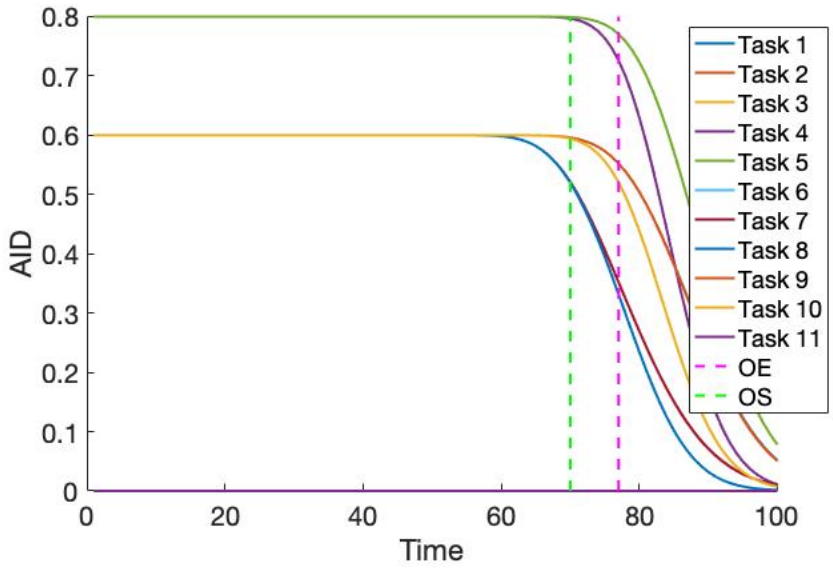


Figure 4.6 AID Value for One Cluster Tool under Different Maintenance Tasks

The time unit is one day, while the operation planning horizon is from Monday to Sunday, i.e., $T = 7$. The starting and ending times of maintenance tasks, i.e., the operation time window, are in the planning horizon for each cluster tool. The duration of each maintenance task is two days. The operation time window of maintenance tasks and the required maintenance tasks (Yao *et al.* 2004) for each cluster tool are defined in Table 4.2. In the table, the required task over a specific cluster tool could start from any date within the earliest and final date node except the last day of the planning horizon, marked in a diamond shape. From the table, one can find that task 1 on cluster tool 1 should be performed between Monday and Wednesday. Task 5 on cluster tool 1 is performed between

Wednesday and Friday, while task 10 is between Friday and Sunday. For cluster tool 8, task 6 must be performed on Monday. For the resource constraint in the case study, we consider the number of operators needed for performing maintenance tasks. The operators needed in a day should not be more than the total number of the available operators.

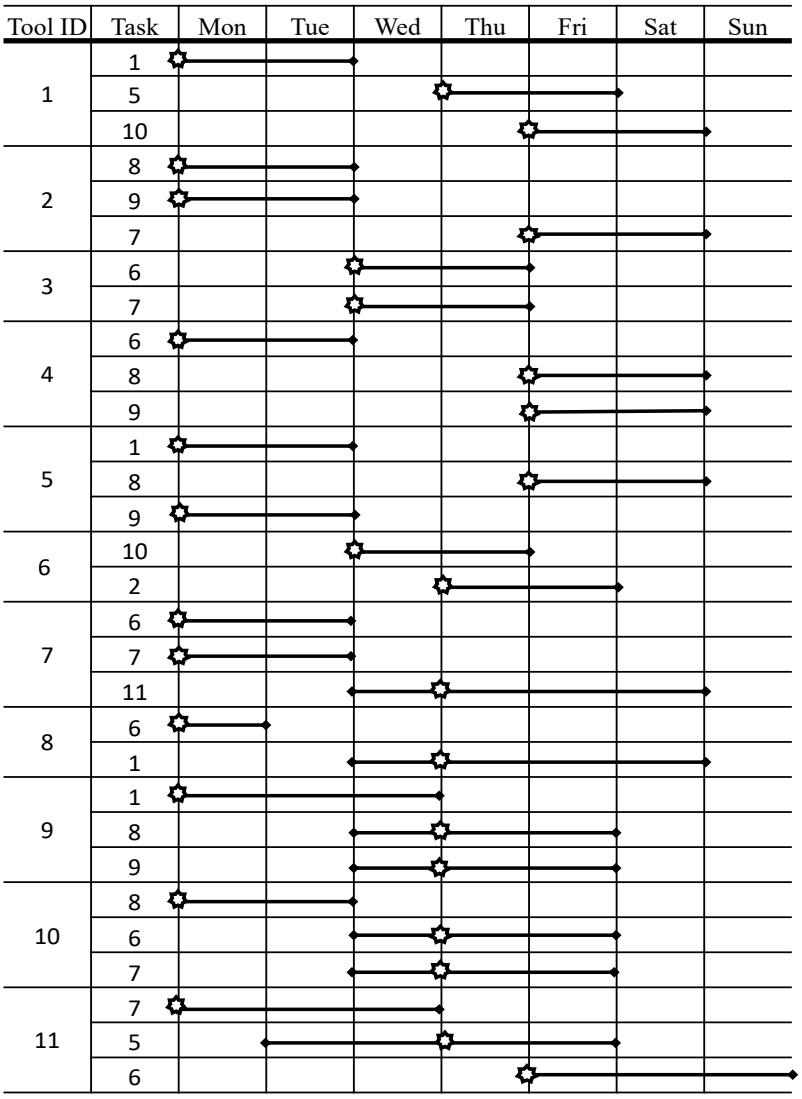
The objective function is now solved under the parameters and constraints presented in section 4.4.3. We use YALMIP (Lofberg, 2004.) and CPLEX to solve this MIP problem. For our simulation case, the MIP model has 480 decision variables and 279 constraints. The solving time is 0.7374 seconds. The optimal maintenance schedules for the cluster tools are shown in Table 4.2 along the scheduling horizon, with tool ID from 1 to 11. The heptagrams represent the optimal maintenance task starting time that is computed by the AID integrated MIP model.

To evaluate the profit improvement brought by the proposed scheduling scheme, we compare it with the random pick-up schedule model and the two-level hierarchical modeling framework (Yao *et al.* 2004). For the random pick-up schedule model, based on the planned operation window, a maintenance task starts and ends within this operation time window. We randomly generated 10000 maintenance schedules according to the operation window. For the two-level hierarchical model, it has eight different PM schedules that achieves the same profit rate. The comparison of the actual profit rate is shown in Figure 4.7. We can see that the proposed AID scheme has a higher profit rate (54.3396) than the randomly picked schedules, which gain the maximum profit rate as 53.34. Our approach is also compared with the two-level hierarchical model (Yao *et al.* 2004), which is the case without AID, and shows significant profit rate improvement.

This simulation study shows that the optimal proactive maintenance scheduling approach outperforms the random pick-up schedules and the two-level hierarchical model

in the sense that the proposed method delivers the optimal schedule with the highest profit rate.

Table 4.2 Maintenance Tasks w.r.t Time Windows (Heptagrams Represents the Optimal Task Starting Time Computed by the AID Integrated MIP Model)



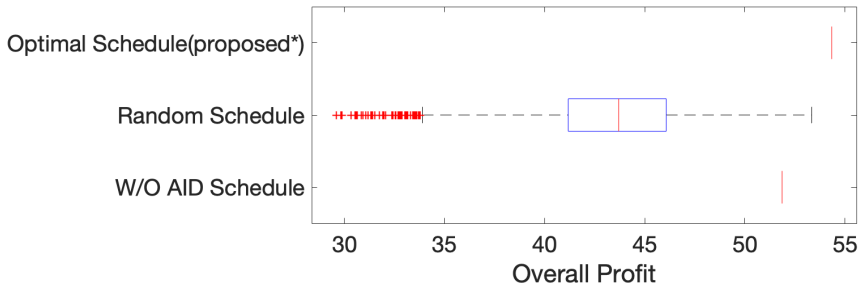


Figure 4.7 Actual Profit Rate of the Proposed New Method Compared with Benchmark Methods

4.6 Conclusion

This chapter develops an AID integrated optimal proactive maintenance scheduling model which integrates prognostic and diagnostics information into a three-level proactive maintenance decision-making framework. The three levels of the model are (1) planning model, (2) Sensor-driven AID scheme, and (3) scheduling model. The planning model determines the operation time window with respect to demand pattern, coordinate the supplement that needed for the maintenance tasks, and empirical status of the cluster tools. The sensor-driven AID scheme provides guidelines to the next level by estimating the data-driven degradation trend of the chambers and integrating the degradation status to the availability of the cluster tools to precisely estimate the expected cluster tools' availability. The scheduling model integrates the AID status of the cluster tools and the planning policies from the planning model and provides an optimal maintenance schedule. This chapter mainly focuses on the sensor-driven AID scheme to integrate sensing signal into prognostics and the scheduling model to use MIP to allocate optimal maintenance schedules with the consideration of the reliability, availability, and resource restrictions.

According to the case study, the proposed approach is feasible and more profitable than the existing maintenance scheduling methods. By applying the proposed framework to surrogate semiconductor industry data, we conclude that the AID integrated optimal proactive maintenance scheduling model can offer an optimized schedule to maintenance decision-making for cluster tools.

Further studies can be done to investigate the degradation status/ health index to achieve improved maintenance strategy. In this chapter, although we provide the formulation and data structure of the degradation model, only a linear degradation model is presented for illustration purposes. However, the degradation trend could be more complicated, and more investigations are needed to obtain more accurate estimation of degradation signals.

CHAPTER 5. SUMMARY AND FUTURE RESEARCH

5.1 Summary of Original Contributions

This dissertation develops methodologies to establish the composite index via data fusion for monitoring, diagnosis, and maintenance decision-making in the various complex multistage manufacturing applications. In this thesis, we integrate advanced statistical methods with engineering domain knowledge to develop novel methodologies. The original contributions of this dissertation include:

- A novel composite index is developed to monitor real-time product quality for the continuous production of CNT buckypaper. The continuous production of CNT buckypaper is a complex multistage manufacturing process, while the Raman spectroscopy sensor is used to collect massive high dimensional signals. The proposed construction scheme for the composite index integrates penalized mixed-effects decomposition (PMD), weighted cross-correlation, and maximum margin clustering methods to deliver a single quality index for fast product quality monitoring. In the case study, the proposed assessment approach distinguishes the quality performance of the different CNT buckypaper samples successfully. It can identify the long-term mean shift that occurs in the process as well as the samples with the large within-sample disorder. It provides quantitative quality indices for single-walled CNT buckypaper after acid processing or functionalization while the quality assessment results are consistent with evaluations of the experienced engineers.

- A novel automatic analytical framework, called "MADE," is proposed to streamline the identification process of the influential process variables related to the quality variable. The main contribution of this study is to build an integrated framework to automatically conduct initial prominent process variable identification for a set of potential quality problems for the hot rolling process. The proposed framework integrates feature extraction, clustering, and automatically identifies the process variables that are influential to the product quality using multiple types of statistical models via multiple permutation tests. The case study uses the hot rolling process in field data to demonstrate that the proposed method can efficiently identify the most influential process variables with minimum involvements of data scientists.
- A composite index, Availability-Degradation (AID), is proposed to integrate the degradation status of sensing information at the component-levels with the machine availability information at the plant-level. We further develop AID integrated optimal proactive maintenance scheduling model, which incorporates prognostic and diagnostics information into a three-level proactive maintenance decision-making framework. The proposed strategy is tested and validated with a simulated semiconductor manufacturing process.

5.2 Future Research

Several important topics can be explored for further development of the composite index and its utilization in the data-rich multi-stage manufacturing system. For example, in the CNT buckypaper process, the developed quality assessment index provides final quality information of the product. In the meantime, the CNT buckypaper fabrication

process is a multi-stage process that has many process variables, include both control factors and noise factors. An automatic root-cause diagnosis and process control algorithm is needed to link the control factors with the final quality variable so that the product quality can be improved eventually.

From the data fusion side, a multi-stage manufacturing process, together with advanced sensing technologies, provides large-size, high-dimensional, and heterogeneous data. Meanwhile, the collected data are usually not well structured due to the characteristics of the multi-stage manufacturing process. Further studies can be done to investigate how to extract meaningful key performance indicators from the massive, heterogeneous data with consideration of unstructured data format.

REFERENCES

Abu-Absi, N.R., Kenty, B.M., Cuellar, M.E., Borys, M.C., Sakhamuri, S., Strachan, D.J., Hausladen, M.C. and Li, Z.J., 2011. Real time monitoring of multiple parameters in mammalian cell culture bioreactors using an in-line Raman spectroscopy probe. *Biotechnology and Bioengineering*, 108(5), pp.1215-1221.

Ajayan, P.M. and Zhou, O.Z., 2001. Applications of Carbon Nanotubes. In *Carbon Nanotubes*. Berlin, Heidelberg: Carbon Nanotubes, pp. 391–425.

Alahbabi, M.N., Cho, Y.T. and Newson, T.P., 2006. Long-range distributed temperature and strain optical fibre sensor based on the coherent detection of spontaneous Brillouin scattering with in-line Raman amplification. *Measurement Science and Technology*, 17(5), pp.1082–1090.

Lofberg, J., 2004, September. YALMIP: A toolbox for modeling and optimization in MATLAB. In *2004 IEEE International Conference on Robotics and Automation* (IEEE Cat. No. 04CH37508), pp. 284-289.

Babbie, E.R., 2015. *The practice of social research.* ISBN-13: 9781305104945

Ben-Daya, M., Kumar, U. and Murthy, D.N.P., 2016. *Introduction to Maintenance Engineering*, John Wiley & Sons.

Best, W.R., Becktel, J.M., Singleton, J.W. and Kern, F., 1976. Development of a Crohn's disease activity index: National Cooperative Crohn's Disease Study. *Gastroenterology*, 70(3), pp.439-444.

Bleakie, A. and Djurdjanovic, D., 2013. Feature extraction, condition monitoring, and fault modeling in semiconductor manufacturing systems. *Computers in Industry*, 64(3), pp.203–213.

Bodis, L., 2007. *Quantification of spectral similarity towards automatic spectra verification.* (Doctoral dissertation, ETH Zurich).

Butte, S., Prashanth, A.R. and Patil, S., 2018, April. Machine learning based predictive maintenance strategy: a super learning approach with deep neural networks. In *2018 IEEE Workshop on Microelectronics and Electron Devices (WMED)* (pp. 1-5). IEEE.

Cha, C., Shin, S.R., Annabi, N., Dokmeci, M.R. and Khademhosseini, A., 2013. Carbon-based nanomaterials: multifunctional materials for biomedical engineering. *ACS nano*, *7*(4), pp.2891-2897.

Chang, T.-S., Shi, J. and Zhou, S., 2009. SQA (TM): Surface Quality Assured Steel Bar Program. (No. DOE/GO/14003-1). OG Technologies, Inc..

Cheng, Q., Bao, J., Park, J., Liang, Z., Zhang, C. and Wang, B., 2009. High mechanical performance composite conductor: multi-walled carbon nanotube sheet/bismaleimide nanocomposites. *Advanced Functional Materials*, 19(20), pp.3219-3225.

Cherry, G.A. and Qin, S.J., 2006. Multiblock principal component analysis based on a combined index for semiconductor fault detection and diagnosis. *IEEE Transactions on Semiconductor Manufacturing*, 19(2), pp.159-172.

De Bondt, W.F. and Thaler, R.H., 1995. Financial decision-making in markets and firms: A behavioral perspective. *Handbooks in Operations Research and Management Science*, 9, pp.385–410.

de Gelder, R., Wehrens, R. and Hageman, J.A., 2001. A generalized expression for the similarity of spectra: application to powder diffraction pattern classification. *Journal of Computational Chemistry*, 22(3), pp.273–289.

Ding, Y., Elsayed, E.A., Kumara, S., Lu, J.C., Niu, F. and Shi, J., 2006. Distributed sensing for quality and productivity improvements. *IEEE Transactions on Automation Science and Engineering*, 3(4), pp.344-359.

Dresselhaus, M.S., Jorio, A., Hofmann, M., Dresselhaus, G. and Saito, R., 2010. Perspectives on carbon nanotubes and graphene Raman spectroscopy. *Nano Letters*, 10(3), pp.751-758.

Eti, M.C., Ogaji, S.O.T. and Probert, S.D., 2006. Reducing the cost of preventive maintenance (PM) through adopting a proactive reliability-focused culture. *Applied Energy*, 83(11), pp.1235–1248.

Fang, X., Gebraeel, N.Z. and Paynabar, K., 2017. Scalable prognostic models for large-scale condition monitoring applications. *IISE Transactions*, 49(7), pp.698–710.

Févotte, G., 2007. In Situ Raman Spectroscopy for In-Line Control of Pharmaceutical Crystallization and Solids Elaboration Processes: A Review. *Chemical Engineering Research and Design*, 85(7), pp.906–920.

Fitch, E.C., 2013. *Proactive Maintenance for Mechanical Systems*, Elsevier Science.

Fox, G.C., Williams, R.D. and Messina, G.C., 2014. Parallel computing works!. Elsevier.

Gebraeel, N.Z. and Lawley, M.A., 2007. A Neural Network Degradation Model for Computing and Updating Residual Life Distributions. *IEEE Transactions on Automation Science and Engineering*, 5(1), pp.154–163.

Gebraeel, N.Z., Lawley, M.A., Li, R. and Ryan, J.K., 2005. Residual-life distributions from component degradation signals: A Bayesian approach. *IIE Transactions,* 37(6), pp.543-557.

Goebel, K. and Bonissone, P., 2005. Prognostic information fusion for constant load systems. *In 2005 7th International Conference on Information Fusion* (Vol. 2, pp. 9-pp). IEEE.

Gommans, H.H., Alldredge, J.W., Tashiro, H., Park, J., Magnuson, J. and Rinzler, A.G., 2000. Fibers of aligned single-walled carbon nanotubes: Polarized Raman spectroscopy. *Journal of Applied Physics*, 88(5), pp.2509-2514.

Gonnet, P., Liang, Z., Choi, E.S., Kadambala, R.S., Zhang, C., Brooks, J.S., Wang, B. and Kramer, L., 2006. Thermal conductivity of magnetically aligned carbon

nanotube buckypapers and nanocomposites. *Current Applied Physics*, 6(1), pp.119-122.

Gunst, R.F., 2018. *Regression analysis and its application: a data-oriented approach. Routledge.* CRC Press.

Guyon, I., Sun-Hosoya, L., Boullé, M., Escalante, H.J., Escalera, S., Liu, Z., Jajetic, D., Ray, B., Saeed, M., Sebag, M. and Statnikov, A., 2019. Analysis of the AutoML challenge series 2015–2018. In *Automated Machine Learning* (pp. 177-219). Springer, Cham.

Heger, T. and Pandit, M.C., 2004. Optical wear assessment system for grinding tools. *Journal of Electronic Imaging*, 13(3), pp.450–462.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, pp.65–70.

Hong, S.J., Lim, W.Y., Cheong, T. and May, G.S., 2011. Fault Detection and Classification in Plasma Etch Equipment for Semiconductor Manufacturing *e*-Diagnostics. *IEEE Transactions on Semiconductor Manufacturing*, 25(1), pp.83-93.

Horne, G. and Liang, Z., 2018. Systems and methods for continuous manufacture of buckypaper materials. U.S. Patent 9, 909, 259.

Jackson, J.E., 2005. A user's guide to principal components. 587. John Wiley & Sons.

Jacques, J. and Preda, C., 2013. Functional data clustering: a survey. *Advances in Data Analysis and Classification*, 8(3), pp.231–255.

Jakubinek, M.B., Kim, K.S., Homenick, C., Kodra, O., Walker, S. and Simard, B., 2019. Assessment of boron nitride nanotube materials using X-ray photoelectron spectroscopy. *Canadian Journal of Chemistry,* 97(6), pp.457-464.

Jalali, A., Heistracher, C., Schindler, A., Haslhofer, B., Nemeth, T., Glawar, R., Sihn, W. and De Boer, P., 2019, June. Predicting Time-to-Failure of Plasma Etching Equipment using Machine Learning. In *2019 IEEE International Conference on Prognostics and Health Management (ICPHM)* (pp. 1-8). IEEE.

Jardine, A.K.S., Lin, D. and Banjevic, D., 2006. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7), pp.1483–1510.

Jin, R., Li, J. and Shi, J., 2007. Quality prediction and control in rolling processes using logistic regression. *Transactions of NAMRI/SME*, 35, pp.113–120.

Jones, R.C., Donaldson, G.C., Chavannes, N.H., Kida, K., Dickson-Spillmann, M., Harding, S., Wedzicha, J.A., Price, D. and Hyland, M.E., 2009. Derivation and validation of a composite index of severity in chronic obstructive pulmonary disease: the DOSE Index. *American Journal of Respiratory and Critical Care Medicine*, 180(12), pp.1189-1195.

Joseph, V.R. and Hung, Y., 2008. Orthogonal-maximin Latin hypercube designs. *Statistica Sinica*, pp.171–186.

Kaihara, T., Fujii, N., Tsujibe, A. and Nonaka, Y., 2010. Proactive maintenance scheduling in a re-entrant flow shop using Lagrangian decomposition coordination method. *CIRP Annals*, 59(1), pp.453-456.

Kaufman, L. and Rousseeuw, P.J., 2009. Finding groups in data: an introduction to cluster analysis. 344. John Wiley & Sons.

Kleiner, A., Talwalkar, A., Sarkar, P. and Jordan, M.I., 2014. A scalable bootstrap for massive data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 76(4), pp.795-816.

Krueger, A., 2010. *Carbon Materials and Nanotechnology*. pp.1–491. John Wiley & Sons.

Lee, W.S., Huang, A.Y., Chang, Y.Y. and Cheng, C.M., 2011. Analysis of decision making factors for equity investment by DEMATEL and Analytic Network Process. *Expert Systems with Applications*, 38(7), pp.8375-8383.

Lehman, J.H., Terrones, M., Mansfield, E., Hurst, K.E. and Meunier, V., 2011. Evaluating the characteristics of multiwall carbon nanotubes. *Carbon*, 49(8), pp.2581-2602.

Leigh, W., Purvis, R. and Ragusa, J.M., 2002. Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support. *Decision Support Systems*, 32(4), pp.361–377.

Li, H., Zhao, N., He, C., Shi, C., Du, X. and Li, J., 2008. Thermogravimetric analysis and TEM characterization of the oxidation and defect sites of carbon nanotubes synthesized by CVD of methane. *Materials Science and Engineering: A*, 473(1-2), pp.355-359.

Li, J., Shi, J. and Chang, T.-S., 2007. On-Line Seam Detection in Rolling Processes Using Snake Projection and Discrete Wavelet Transform. *Journal of Manufacturing Science and Engineering*, 129(5), pp.926–933.

Liu, K., Chehade, A. and Song, C., 2017. Optimize the Signal Quality of the Composite Health Index via Data Fusion for Degradation Modeling and Prognostic Analysis. *IEEE Transactions on Automation Science and Engineering*, 14(3), pp.1504–1514.

Liu, K., Gebraeel, N.Z. and Shi, J., 2013. A Data-Level Fusion Model for Developing Composite Health Indices for Degradation Modeling and Prognostic Analysis. *IEEE Transactions on Automation Science and Engineering*, 10(3), pp.652–664.

Liu, K., Jain, S. and Shi, J., 2012. Physician performance assessment using a composite quality index. *Statistics in Medicine*, 32(15), pp.2661–2680.

Liu, T. and Kumar, S., 2003. Quantitative characterization of SWNT orientation by polarized Raman spectroscopy. *Chemical Physics Letters*, 378(3-4), pp.257–262.

Lu, C.J. and Meeker, W.O., 2012. Using Degradation Measures to Estimate a Time-to-Failure Distribution. *Technometrics*, 35(2), pp.161–174.

Luo, M., Yan, H.C., Hu, B., Zhou, J.H. and Pang, C.K., 2015. A data-driven two-stage maintenance framework for degradation prediction in semiconductor manufacturing industries. *Computers & Industrial Engineering*, 85, pp.414-422.

Matthews, B.W., 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure*, 405(2), pp.442–451.

Menardi, G. and Torelli, N., 2014. Training and assessing classification rules with imbalanced data. *Data Mining and Knowledge Discovery*, 28(1), pp.92–122.

Moyne, J., Samantaray, J. and Armacost, M., 2016. Big data capabilities applied to semiconductor manufacturing advanced process control. *IEEE Transactions on Semiconductor Manufacturing*, 29(4), pp.283-291.

Muller, A., Suhner, M.-C. and Iung, B., 2008. Formalisation of a new prognosis model for supporting proactive maintenance implementation on industrial system. *Reliability Engineering and System Safety*, 93(2), pp.234–253.

Munirathinam, S. and Ramadoss, B., 2014. Big data predictive analtyics for proactive semiconductor equipment maintenance. In *2014 IEEE International Conference on Big Data (Big Data)*. IEEE, pp. 893–902.

Murray, C.J., 1994. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bulletin of the World health Organization*, 72(3), p.429.

Jin, N., Zhou, S., Chang, T.S. and Huang, H.H.H., 2008. Identification of influential functional process variables for surface quality control in hot rolling processes. *IEEE Transactions on Automation Science and Engineering*, 5(3), pp.557-562.

Osswald, S., Havel, M. and Gogotsi, Y., 2007. Monitoring oxidation of multiwalled carbon nanotubes by Raman spectroscopy. *Journal of Raman Spectroscopy*, 38(6), pp.728–736.

Pan, E., Ye, L., Shi, J. and Chang, T.S., 2009. On-line bleeds detection in continuous casting processes using engineering-driven rule-based algorithm. *Journal of Manufacturing Science and Engineering*, 131(6).

Park, J.G., Li, S., Liang, R., Zhang, C. and Wang, B., 2008. Structural changes and Raman analysis of single-walled carbon nanotube buckypaper after high current density induced burning. *Carbon,* 46(9), pp.1175-1183.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J., 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct), pp.2825-2830.

Perkinson, T.L., McLarty, P.K., Gyurcsik, R.S. and Cavin, R.K., 1994. Single-wafer cluster tool performance: An analysis of throughput. *IEEE Transactions on Semiconductor Manufacturing*, 7(3), pp.369-373.

Powers, D.M., 2011. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2, pp. 37–63.

Puterman, M.L., 2014. *Markov Decision Processes*, John Wiley & Sons.

Raravikar, N.R., Keblinski, P., Rao, A.M., Dresselhaus, M.S., Schadler, L.S. and Ajayan, P.M., 2002. Temperature dependence of radial breathing mode Raman frequency of single-walled carbon nanotubes. *Physical Review B,* 66(23), p.235424.

Rokach, L. and Maimon, O., 2005. Clustering Methods. In *Data Mining and Knowledge Discovery Handbook*. New York: Springer, Boston, MA, pp. 321–352.

Nardo, M., Saisana, M., Saltelli, A. and Tarantola, S., 2005. Tools for composite indicators building. *European Comission, Ispra*, 15(1), pp.19-20.

Scarf, P.A., 2003. On the application of mathematical models in maintenance. *European Journal of Operational Research*, pp.1–14.

Shi, J. and Zhou, S., 2009. Quality control and improvement for multistage systems: A survey. *IIE Transactions*, 41(9), pp.744–753.

Sun, Q., 2002, July. Sensor fusion for vehicle health monitoring and degradation detection. In *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002*. (IEEE Cat. No. 02EX5997) (Vol. 2, pp. 1422-1427). IEEE.

Susto, G.A., Conference, A.B.2.I.2.I. 2016, Dealing with time-series data in predictive maintenance problems. *IEEE Transactions on Automation Science and Engineering*.

Trigueiro, J.P.C., Silva, G.G., Lavall, R.L., Furtado, C.A., Oliveira, S., Ferlauto, A.S., Lacerda, R.G., Ladeira, L.O., Liu, J.W., Frost, R.L. and George, G.A., 2007. Purity evaluation of carbon nanotube materials by thermogravimetric, TEM, and SEM methods. *Journal of Nanoscience and Nanotechnology*, 7(10), pp.3477-3486.

Volponi, A., Brotherton, T. and Luppold, R., 2004, September. Development of an information fusion system for engine diagnostics and health management. In *AIAA 1st Intelligent Systems Technical Conference* (p. 6461).

Xu, G., Zhang, Q., Zhou, W., Huang, J. and Wei, F., 2008. The feasibility of producing MWCNT paper and strong MWCNT film from VACNT array. *Applied Physics A,* 92(3), pp.531-539.

Xu, L., Neufeld, J., Larson, B. and Schuurmans, D., 2005. Maximum margin clustering. *In Advances in Neural Information Processing Systems,* pp. 1537-1544.

Yao, X., Fernández-Gaucherand, E., Fu, M.C. and Marcus, S.I., 2004. Optimal preventive maintenance scheduling in semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing,* 17(3), pp.345-356.

Yildirim, M., Sun, X.A. and Gebraeel, N.Z., 2016a. Sensor-driven condition-based generator maintenance scheduling—Part I: Maintenance problem. *IEEE Transactions on Power Systems*, 31(6), pp.4253–4262.

Yildirim, M., Sun, X.A. and Gebraeel, N.Z., 2016b. Sensor-driven condition-based generator maintenance scheduling—part II: incorporating operations. *IEEE Transactions on Power Systems*, 31(6), pp.4263–4271.

Yue, X., Yan, H., Park, J.G., Liang, Z. and Shi, J., 2017. A Wavelet-Based Penalized Mixed-Effects Decomposition for Multichannel Profile Detection of In-Line Raman Spectroscopy. *IEEE Transactions on Automation Science and Engineering*, 15(3), pp.1258-1271.

Yue, X., Wang, K., Yan, H., Park, J.G., Liang, Z., Zhang, C., Wang, B. and Shi, J., 2016. Generalized wavelet shrinkage of inline Raman spectroscopy for quality monitoring of continuous manufacturing of carbon nanotube buckypaper. *IEEE Transactions on Automation Science and Engineering,* 14(1), pp.196-207.

Yue, X., Park, J.G., Liang, Z. and Shi, J., 2020. Tensor Mixed Effects Model with Application to Nanomanufacturing Inspection. *Technometrics,* 62(1), pp.116-129.

Zhang, C., Yan, H., Lee, S. and Shi, J., 2018. Multiple profiles sensor-based monitoring and anomaly detection. *Journal of Quality Technology*, 50(4), pp.344-362.

Zhang, K., Tsang, I.W. and Kwok, J.T., 2009. Maximum margin clustering made practical. *IEEE Transactions on Neural Networks*, 20(4), pp.583-596.

Zhou, R.R., Serban, N. and Gebraeel, N., 2011. Degradation modeling applied to residual lifetime prediction using functional data analysis. *The Annals of Applied Statistics*, 5(2B), pp.1586–1610.